# INTERAURAL CROSS CORRELATION: SOURCES OF VARIABILITY IN CONCERT HALLS

By

GARY S. MADARAS

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

1996

# ACKNOWLEDGMENTS

I want to thank each of my committee members for their guidance. Dr. Earl M. Starnes (Chairperson) (Professor Emeritus, Department of Architecture) was the director of doctoral studies at the onset of the research. Despite his retirement, he cared enough to continue his involvement. Dr. Starnes also helped to make the transition into doctoral studies surprisingly easy.

Gary W. Siebein (Cochairperson) (Professor, Department of Architecture) provided over four years of continual guidance, support, and friendship. There are no words meaningful enough to express the true gratitude that he deserves. He is one of only four people in my life that have truly raised my academic, professional, and personal standards.

Dr. David M. Green (Graduate Research Professor, Psychology), Willis R. Bodine Jr. (Professor, Music), and Bertram Y. Kinzey Jr. (Professor Emeritus, Department of Architecture) completed the supervisory committee. Their interdisciplinary contributions helped to make the research project and my education in general more well-rounded and complete.

I would like to thank the following students for their ideas, encouragement, and friendship: Avraham Bortnick, Rich Cervone, Wei-Hwa Chiang, Martin Gold, Chris Herr, John Kidwell, Rob Lilkendey, Loren Raia, and Mitchell Spolan. Dr. Harold Doddington and Dr. Bill Schwab (Department of Aerospace Engineering, Mechanics, and Engineering Sciences) developed much of the measurement instrumentation and analysis software.

Mr. Art Alvarez (Visual Arts Director, Belk Lindsey, Gainesville) donated a department store manikin for modifications and use during data collection. Dr. Everett Scroggie (Consulting Clinical Audiologist, Gainesville) donated his time and advice, making the design of the full scale manikin accurate. Dr. Glenn E. Turner and Mr. Lee D. Mintz in the Department of Prosthodontics (College of Dentistry, University of Florida) generously donated their time and materials to cast silicon pinnae for the full scale manikin. Dr. John C. Middlebrooks in the Departments of Neuroscience and Surgery (University of Florida) measured the full scale manikin's head related transfer functions, supplied human subject data for comparison, and discussed the results of the data analysis.

The Concert Hall Research Group, an organization of professional acoustical consulting firms, researchers, and scholars, funded and supported a trip to collect acoustical data in multiple Northeastern United States concert halls. The data collected during that trip were also used for this research project. Drawings of the concert halls which were used for reference during this research project were copied from originals that were prepared for and funded by the Concert Hall Research Group.

Dr. John Bradley (National Research Council of Canada) generously gave access to data collected during the Concert Hall Research Group trip. The ability to analyze an additional set of data was beneficial and helped to establish even further confidence in the results and conclusions.

Jaffe Holden Scarbrough Acoustics Inc. (Norwalk, CT) generously provided access to equipment for the physical document preparation, printing, and copying.

Finally, I want to thank my family for believing in me, and providing me with love and support.

## PREFACE

When listening to music some listeners may prefer an experience that makes them feel close to the performance (intimacy) with the ability to distinguish each note in the staccato passages (clarity). Others may prefer a different experience, perhaps one that reinforces the natural sound from the performers (loudness) and causes it to persist throughout the room decaying slowly (reverberance). The relative importance of these preferences varies from one listener to another, one musical style to another, and perhaps even from one performance to another.

Researchers in the field of architectural acoustics attempt to relate listener preference and the architectural design of enclosures. One commonly used method of relating preference to quantitative aspects of architectural design such as room volume or seating area involves an intermediate step, one that tries to represent the behavior of sound energy within an enclosure with quantitative parameters. This approach requires a dual understanding. The first part is to know how listener preference for an actual performance can be predicted using quantitative acoustic parameters. The second part is to know how these parameters are influenced by the architectural design of the enclosure. This research investigates the sources of variability of one particular acoustic parameter, interaural cross correlation, and identifies how this measure can be influenced by architectural design.

TABLE OF CONTENTS

## LIST OF TABLES

LIST OF FIGURES

Abstract of Dissertation Presented to the Graduate School
of the University of Florida in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy

INTERAURAL CROSS CORRELATION:
SOURCES OF VARIABILITY IN CONCERT HALLS

By

Gary S. Madaras

May, 1996

Chairperson: Dr. Earl Starnes
Cochairperson: Gary Siebein
Major Department: Architecture

After constructing a binaural hearing manikin, monaural and binaural impulse responses were collected inside multiple concert halls. Data analysis identified the level of the direct sound relative to the level of the architectural reflections (calculated using a monaural impulse response) as a significant source of variability between interaural cross correlation values. Excluding the direct sound from the interaural cross correlation integral durations significantly decreased the variability between the calculated values. The variability that remained seemed to relate to the arrival direction of the architectural reflections.

To further investigate the effect of reflection arrival direction on interaural cross correlation, a measurement method for scale models was developed. Different model configurations with various architectural elements and surface treatments were tested. Model configurations that had reflections approaching the receiver from the sides produced much

lower interaural cross correlation values than those that did not have reflections approaching the receiver from the sides. The various configurations of architectural elements (*i.e.*, balconies and stage canopy) and surface treatments (*i.e.*, diffusive and specular) affected interaural cross correlation values to a lesser extent than altering the arrival direction of the architectural reflections.

It was concluded that the general direction from which the architectural reflections arrive and the finish of the architectural surfaces are both sources of variability between interaural cross correlation values. The basic direction from which most of the architectural reflections arrive affects interaural cross correlation to a much greater extent than either the placement of smaller architectural elements within the room or the finish of the architectural surfaces. However, the arrival direction of the architectural reflections and the finish of the architectural surfaces must both be considered in order to achieve the lowest interaural cross correlation values. This means that early schematic design decisions involving the orientation and proximity of primary reflecting surfaces relative to the audience (*i.e.*, room shape) can only affect interaural cross correlation values to a certain extent. To achieve even lower values, the placement of smaller architectural elements within the room as well as the finish materials applied to the architectural surfaces must also be considered.

# INTERAURAL CROSS CORRELATION

The basic understanding of reverberation has been known for centuries. A precise reverberation time formula relating two architectural sources of variability, room volume and the amount of absorption, was developed by Wallace C. Sabine over ninety years ago. In contrast, using two acoustic signals recorded from the left and right ears of a directional receiver to gain information beyond that which can be obtained from a single signal recorded omni-directionally is a fairly recent development. It was in the 1960s that architectural acousticians began comparing the signals from a binaural receiver, hoping to derive a criterion that indicated the primary arrival direction of the architectural reflections while also relating to human preference for music listening. Proposed binaural criteria have been based on the direct multiplication of the two signals followed by integration over various durations.

Theoretically, integrating the product of the two signals should give some indication of the arrival direction of the architectural reflections. Sound approaching the binaural receiver from incidence angles within the median plane of the head would result in equally high sound pressures in both ears and consequently large products when the two signals are multiplied. After integration, an overall high value would result. This condition typically occurs in fan-shaped rooms as a result of overhead reflections off the low ceiling planes. Conversely, sound approaching from angles outside the median plane of the head would result in the shadowed ear experiencing lower sound pressure than the exposed ear. The product of the two signals would be smaller, and after integration the resulting value would be lower. This condition typically occurs in rectangular rooms as a result of reflections off the side walls.

This explanation assumes that the total integrated energy reaching the binaural receiver remains constant, and that only the integrated product of the two signals changes. Once the total integrated energy varies, a change in the integrated product of the left and right signals could be due to a change in the arrival direction of the architectural reflections or a change in the overall level independent of direction. In other words, the integrated product of two loud signals, even if they are greatly different from each other (indicating incidence angles outside the median plane of the head), could be greater than that of two similar signals having extremely low levels. In this case, the variability due to the overall level difference obscures any directional information.

In order to gain consistently reliable directional information, the comparison of the left and right signals needs to be normalized so that the overall level of the sound is not a source of variability. In 1968, Danilenko (as part of his graduate research in Aachen) defined a binaural distinctness coefficient using equation 1, which attempts to normalize for loudness by using a ratio rather than an absolute value (Cremer and Mueller, 1982). It is basically a binaural early/total energy ratio resembling its monaural counterpart for distinctness (Thiele, 1953).

Equation 1

$$\zeta_B(t_2) = \frac{\int_{t_1}^{t_2} P_{lt}(t) P_{rt}(t)\,dt}{\int_{t_1}^{\infty} P_{lt}(t) P_{rt}(t)\,dt}$$

$\zeta_B$ Binaural distinctness coefficient.
$P_{lt}$ Sound pressure in the left ear at time $(t)$.
$P_{rt}$ Sound pressure in the right ear at time $(t)$.
$t_1$ This value is 0.0ms, the instant when the direct sound reaches the receiver.
$t_2$ Danilenko fixed this value at 50ms as borrowed from Thiele's (1953) definition of Deutlichkeit.

Despite Danilenko's attempt to normalize the parameter against overall loudness, two factors unrelated to the arrival direction of the architectural reflections influenced the calculated values. The first factor was that both the numerator and the denominator were dependent on the arrival direction of the architectural reflections. Therefore, the early to total binaural energy ratio of two similar signals could equal the same ratio of two greatly different signals. The second factor was that both the numerator and the denominator could assume positive or negative values. As a result, similar signals could produce a surprisingly low value simply due to a negative sign in lieu of a positive one. Succeeding researchers attempted to decrease this variability by eliminating the possibility of negative values, and offering a denominator that is not dependent on the direction of the incoming reflections.

Cremer and Mueller (1982) suggested using the cross correlation function in the denominator of the equation as a reliable normalizer, for it is not influenced by the arrival direction of the architectural reflections. Their interaural cross correlation formula is shown in equation 2.

Equation 2

$$K_{(\tau)} = \frac{\int_{t_1}^{\infty} P_{lt}(t) P_{rt}(t+\tau)\,dt}{\left(\int_{t_1}^{\infty} P_{lt}^2(t)\,dt \int_{t_1}^{\infty} P_{rt}^2(t)\,dt\right)^{1/2}}$$

$K$ Interaural cross correlation as a function of *tau*.
$P_{lt}$ Sound pressure in the left ear at time ($t$).
$P_{rt}$ Sound pressure in the right ear at time ($t$).
$t_1$ This value is 0.0ms, the instant when the direct sound reaches the binaural receiver.
*tau* Amount of time that the right signal is shifted relative to the left signal before the calculation is performed.

In addition to the cross correlation function as a normalizer, Cremer and Mueller also introduced a time shift of the right signal relative to the left signal designated by *tau*. The result of Cremer and Mueller's equation is not a single number, but instead interaural cross correlation as a function of *tau*. The time shift provided a method of compensation in the event that the incidence angle of the direct sound was not normal to the interaural axis of the binaural receiver during the measurement session. If the receiver was oriented so that the direct sound reached the two ears simultaneously, the interaural cross correlation value at *tau* = 0 was used. However, if the receiver was rotated slightly relative to the incidence angle of the source, the interaural cross correlation value at *tau* equal to the resulting interaural time delay was used.

Since *tau* was meant to compensate for only slight head angles relative to the direct sound, the time shift need only be in the order of 500 microseconds or less. However, since the direct sound of the right signal can either precede or succeed that in the left, *tau* needs to assume both positive and negative values. A total *tau* range of 1ms (-500 microseconds to +500 microseconds) has often been interpreted to be a *tau* range of -1ms to +1ms (a total range that exceeds that which is necessitated by a factor of two). Damaske (1968) realized that the maximum value along Cremer and Mueller's *tau* versus interaural cross correlation curve typically occurred when the direct sounds to the two ears were aligned. Therefore, only the maximum value of $K_{(tau)}$ needed to be calculated.

Cremer and Mueller themselves recognized faults in the logic of including the time shift in the interaural cross correlation function. A specific example is when the direct sound to both ears is attenuated from grazing over seat backs or even blocked by architectural surfaces such as balcony fasciae. The succeeding architectural reflections, especially in balcony seats, are commonly louder than the direct sound. In these instances the maximum along the *tau* curve

would designate the alignment of some undetermined reflection instead of the direct sound. As a result, two signals that were initially aligned could be skewed by the *tau* shift.

One must question the logic of using the *tau* shift even when used as it was originally intended (*i.e.*, to align the direct sounds of the two signals). For example, assume that the binaural receiver was angled slightly off axis relative to the incidence angle of the direct sound, and as a result a slight interaural time difference between the direct sounds to the ears occurred. Even though the direct sounds may not be aligned (a condition that commonly occurs when an audience member is listening to a performance), the interaural relationship of all succeeding reflections is true. If the *tau* shift is then used to align the direct sounds, it begins to compensate for measurement method inconsistencies (namely receiver orientation relative to the source) but it also misaligns the numerous succeeding architectural reflections. The theoretical validity of using the *tau* shift becomes more questionable as the interaural time difference between the direct sounds increases.

It would be interesting to study the effect on listener preference as the two signals are shifted out of their original alignment (*tau* = 0) to where the maximum interaural cross correlation value occurs, especially for balcony seats or positions where the direct sound is blocked. In these two cases, the maximum interaural cross correlation value would correspond to the alignment of some undefined reflection instead of the direct sounds. One can speculate that in preference tests, a sound field with the left and right signals shifted by a large *tau* value would be rated quite differently, and perhaps much lower, than a sound field with its original alignment between the left and right signals preserved (*tau* = 0).

Short of removing the *tau* shift from the commonly used interaural cross correlation formula, researchers that question its use can simply take special care to consistently orient the interaural axis of the binaural receiver so that it is perpendicular to the incidence angle of the

direct sound. If this is practiced, the maximum interaural cross correlation value should occur at *tau* = 0 (refer to Ando, 1977a and 1985).

Positioning the receiver towards the source is merely one aspect of establishing a consistent measurement method. It alleviates any concern that may result from altering the original alignment of the two signals, since the effect of the *tau* shift on listening preference remains unknown. A receiver orientation toward the source is only intended to represent the many head orientations held by actual listeners throughout the duration of a performance.

Recognizing some of the questions regarding the use of the *tau* shift, Keet (as part of his graduate research at Capetown University) proposed a slight variation to Cremer and Mueller's interaural cross correlation formula. Keet fixed *tau* equal to 0.0, and used much shorter integral durations. He believed that Cremer and Mueller's longer integral durations related more to the overall diffuseness of the sound field since it was proportionally dominated by the later arriving reverberant energy as opposed to the early directional reflections arriving soon after the direct sound. Since Keet was less interested in total diffuseness and more interested in distinguishing the amount of early sound approaching from the sides relative to that which was arriving from the front, he limited $t_2$ in equation 3 to 50ms.

Equation 3

$$K_{(\tau=0)} = \frac{\int_{t_1}^{t_2} P_{lt}(t) P_{rt}(t+\tau)\,dt}{\left(\int_{t_1}^{t_2} P_{lt}^2(t)\,dt \int_{t_1}^{t_2} P_{rt}^2(t)\,dt\right)^{1/2}}$$

$K$ Interaural cross correlation for *tau* = 0.
$P_{lt}$ Sound pressure in the left ear at time ($t$).
$P_{rt}$ Sound pressure in the right ear at time ($t$).
$t_1$ This value is 0.0ms, the instant when the direct sound reaches the binaural receiver.
$t_2$ This value is 50ms, including early reflections but not the later reverberant energy.
*tau* Amount of time that the right signal is shifted relative to the left signal before the calculation is performed. Keet set *tau* equal to 0.0.

In 1973, Gottlob (as part of his graduate work at Gottingen) combined Keet's shorter integral duration (*i.e.*, $t_2$ = 50ms) with Damaske's proposal to use the maximum value along Cremer and Mueller's *tau* function to further refine the most commonly used formula for calculating interaural cross correlation (refer to equation 4).

Equation 4

$$IACC_{(\max)} = \frac{\int_{t_1}^{t_2} P_{lt}(t) P_{rt}(t+\tau)\,dt}{\left(\int_{t_1}^{t_2} P_{lt}^2\,dt \int_{t_1}^{t_2} P_{rt}^2\,dt\right)^{1/2}}$$

IACC Maximum interaural cross correlation value.
$P_{lt}$ Sound pressure in the left ear at time *(t)*.
$P_{rt}$ Sound pressure in the right ear at time ($t$).
$t_1$ This value is 0.0ms, the instant when the direct sound reaches the binaural receiver.
$t_2$ Values of 50ms, 80ms, and 100ms are common. These integral durations are considered to be early interaural cross correlation and have been found to relate qualitatively to spaciousness and perceived source width (Soulodre and Bradley, 1994). Some researchers (Hidaka *et al.*, 1991) also suggest a late interaural cross correlation where $t_2$ is as long as 3.0sec. Late interaural cross correlation has been related to the diffusive qualities of rooms (envelopment).
*tau* Amount of time that the right signal is shifted relative to the left signal before interaural cross correlation is calculated.

Despite differing opinions regarding the actual value of $t_2$, equation 4 has remained the most common method of calculating interaural cross correlation from binaural impulse responses. Even though interaural cross correlation is a recently developed measure, especially relative to reverberation time, the published literature consistently shows that interaural cross correlation values are strongly related to preference for music listening.

# RELATIONSHIP BETWEEN INTERAURAL CROSS CORRELATION AND LISTENING PREFERENCE

## Gottingen Studies

### Introduction

The first decade of binaural research in architectural acoustics primarily examined different methods of comparing the left and right ear signals. It was not until the Gottingen studies in the early 1970s that Schroeder, Gottlob, and Siebrasse (1974) attempted to relate subjective preference for music listening to multiple quantitative measures, including interaural cross correlation. A review of their methods and conclusions will show how the Gottingen studies first related lower interaural cross correlation values with preferred listening.

### Method

A two track recording of Mozart's Jupiter Symphony (recorded by the English Chamber Orchestra in anechoic conditions) was played through two nondirectional loudspeakers spatially separated on the stages of twenty-two European concert halls. The music was re-recorded in each of the halls using a binaural manikin. Playback occurred through two loudspeakers in front of listeners seated inside an anechoic chamber. The loudspeakers were separated from each other, and steps were taken to prevent cross-talk (*i.e.*, the sound from one speaker did not cancel important information from the other). Listeners were able to toggle back and forth between two halls at a time and even request replay. Twelve subjects were asked to perform preference tests whereby the preferred hall received a score of +1 and the other received a score of -1. If no preference was established, both halls received a score of 0.

Results

Qualitative judgments were correlated with six quantitative parameters including hall volume, hall width, initial time delay gap, reverberation time, definition, and interaural cross correlation. Figures 1 and 2 show the results of the Gottingen studies. Figure 1 shows the results for the eleven halls with reverberation times below 2.0 seconds. When the vertical component of the interaural cross correlation vector (C) is projected down to the consensus preference axis (D1) a correlation coefficient of approximately -0.60 is found. This indicated that as interaural cross correlation decreased, subject preference increased. Figure 2 shows the results of the eleven halls with reverberation times greater than 2.0 seconds. A similar projection technique yields an even higher correlation coefficient of approximately -0.75.

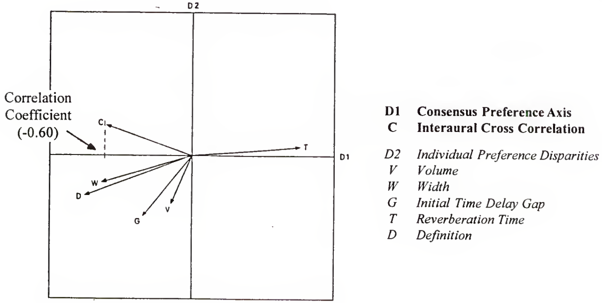



Figure 1 Quantitative parameters correlated with preference (Schroeder *et al.*, 1974)

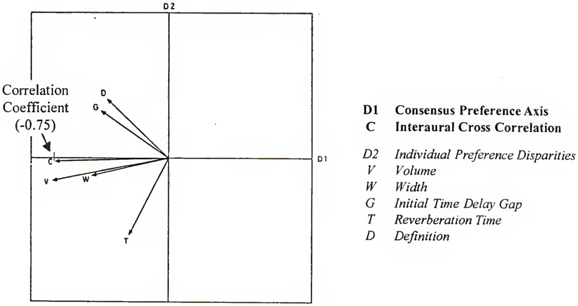



Figure 2 Quantitative parameters correlated with preference (Schroeder *et al.*, 1974)

<u>Conclusions</u>

The researchers concluded that interaural cross correlation is significantly related to preference. It was found to be one of the most important factors for music listening. Just as important was the fact that interaural cross correlation was uncorrelated with reverberation time, meaning that interaural cross correlation has an independent subjective significance. The Gottingen researchers related lower interaural cross correlation values to the feeling of being immersed in the sound. The authors also stated that a method of architecturally achieving lower interaural cross correlation values in concert halls had yet to be defined. They suggested a possible relationship to diffusers, but stated that further study was merited.

# Ando's Studies

## Introduction

During the mid to late 1970s, Ando performed a series of experiments using the facilities in Gottingen (soon after Gottlob and Siebrasse finished their graduate work). The primary purpose of Ando's initial experiments (1977a, 1977b, 1979a, 1979b) was to relate both initial time delay gap and interaural cross correlation to preference for music listening. His results showed that preference increased as interaural cross correlation decreased.

Ando later continued his studies, with the goal of being able to predict subjective preference in concert halls using optimum design objectives. The results of Ando's cumulative efforts (1975 through 1984) are presented in his book Concert Hall Acoustics (1985). In this book, Ando clearly states that interaural cross correlation is one of only four quantitative parameters needed to predict listener preference in concert halls. The studies performed by Schroeder, Gottlob, and Siebrasse as well as those completed by Ando form the empirical foundation which supports the correlation between listening preference and lower interaural cross correlation values.

## Experiment 1

Method. The first experiment (Ando, 1977a) used a computer to simulate the direct sound (azimuth 0, elevation +9) and a single architectural reflection (varying in azimuth from 0 to +90 in 18 degree increments and having a fixed delay of 32ms)[1]. The source signals, two

---

[1] Ando defined incidence angles for the direct sound and subsequent reflections with a double-pole coordinate system (Knudsen, 1982). The horizontal position is noted as the azimuth angle so that 0 degrees is forward, +90 degrees is to the side of the microphone, and -90 degrees is to the opposite side of the microphone. The vertical location is noted as the elevation angle and is measured relative to a horizontal plane. Positions below the horizontal plane containing the subject's interaural axis are denoted with negative values while those above are denoted with positive values.

different music motifs, were played to each of 15 subjects inside an anechoic chamber using loudspeakers. Paired comparison tests were performed by giving the preferred sound field a score of +1 and the other a score of -1. No preference resulted in a score of 0.

Results. Figure 3 plots preference and interaural cross correlation versus azimuth angle for both music motifs (*A* and *B*). Azimuth angles close to 0 degrees resulted in low preference and high interaural cross correlation values. As the azimuth angle increased towards 60 degrees, interaural cross correlation decreased and preference increased. For azimuth angles between 60 degrees and 90 degrees, preference generally decreased and interaural cross correlation generally increased, but the relationship is not as strong as that for azimuth angles less than 60 degrees. The correlation coefficient between preference and interaural cross correlation was -0.76 (1% significance level).

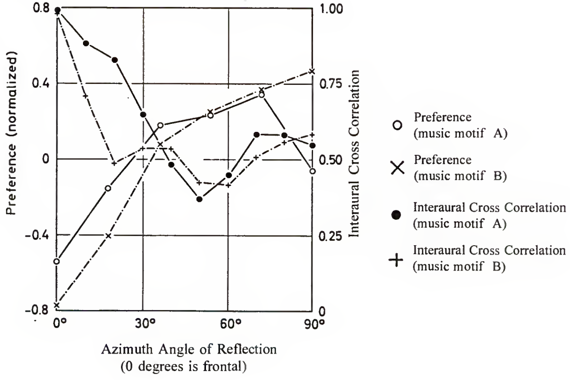



Figure 3 Preference and interaural cross correlation versus reflection azimuth angle (Ando, 1977a)

Conclusions. Ando concluded at the end of his first experiment that the preference score of the sound field increased by decreasing interaural cross correlation, so that the preferred echo directions were found in a range centered on 55 degrees.

Experiment 2

Method. Ando's second experiment (1977b) was similar to the first except for three differences. First, the simulated direct sound originated from a 0 degree azimuth angle and a 0 degree elevation angle. Second, the single architectural reflection varied in azimuth angle from 0 degrees to 180 degrees in 15 degree increments and had a fixed delay of 16ms. Third, speech was used as the source instead of music.

Results. Preference and interaural cross correlation are plotted against the azimuth angle of the simulated architectural reflection in figure 4. As in the results of the first experiment, preference and interaural cross correlation are negatively correlated. As the azimuth angle of

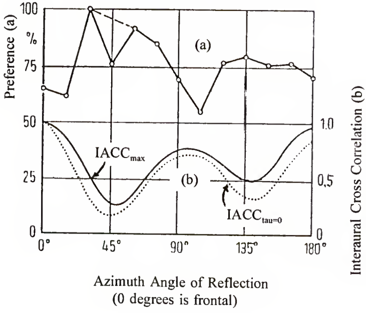


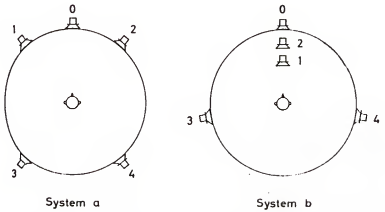

Figure 4 Preference and interaural cross correlation versus reflection azimuth angle (Ando, 1977b)

the simulated architectural reflection increases from 0 degrees to 45 degrees, preference generally increases and interaural cross correlation decreases. The correlation coefficient between preference and interaural cross correlation was -0.71 (1% significance level).

Conclusions. Ando again concluded that lower interaural cross correlation values resulted in better preference ratings. The results were similar to those found by Schroeder, Gottlob, and Siebrasse. They also found that the interaural cross correlation is a significant parameter.

Experiment 3

Method. Ando's third experiment (1979a) simulated the direct sound (azimuth 0, elevation +9) and four early architectural reflections with two different speaker systems (refer to figure 5 for diagrams and arrival directions). In addition, each spatial system had two different level patterns (refer to figure 6 for levels and arrival times). Once again, approximately fifteen subjects performed paired comparison tests while seated in an anechoic chamber.




Arrival Direction (degrees)

| | Azimuth | Elevation |
|---|---|---|
| 0 | 0 | +9 |
| 1 | -40 | +27 |
| 2 | +40 | +27 |
| 3 | -140 | +27 |
| 4 | +140 | +27 |

Arrival Direction (degrees)

| | Azimuth | Elevation |
|---|---|---|
| 0 | 0 | +9 |
| 1 | 0 | +45 |
| 2 | 0 | +27 |
| 3 | -100 | +27 |
| 4 | +100 | +27 |

Figure 5 Speaker systems *a* and *b* (Ando, 1979a)

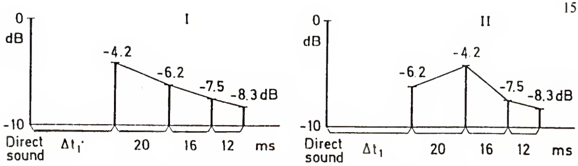



Figure 6 Level and time patterns *I* and *II* (Ando, 1979a)

Results. Figure 7 shows the results of the preference tests for the two speakers systems (*a* and *b*), two level patterns (*I* and *II*), and two music motifs (*A* and *B*). Ando found that interaural cross correlation depended only slightly on level pattern and music motif, but differed significantly for the two speaker systems. Figure 7 shows that the preference for speaker system *a* (interaural cross correlation = 0.53) was higher than that for speaker system *b* (interaural cross correlation = 0.74 to 0.88)

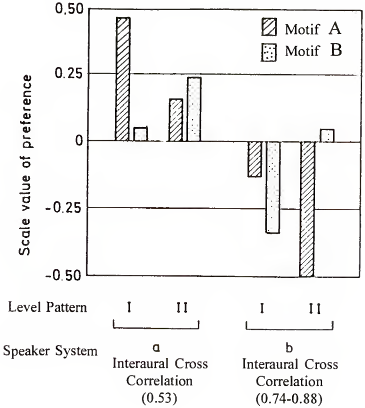



Figure 7 Preference related to interaural cross correlation (Ando, 1979a)

Conclusions. Ando concluded that the preference tests for sound fields with multiple reflections gave nearly the same results as the previous experiments with only one reflection. (*i.e.*, The preference scores of sound fields having the same temporal pattern of reflections decrease with increased interaural cross correlation.)

Experiment 4

Method. Ando's fourth experiment (1979b) simulated the direct sound, multiple early reflections with varying patterns, and a subsequent reverberant field. The two speaker systems shown in figure 5 were again used during this experiment. Four different music motifs were used. The temporal and level pattern of the simulated sound field is shown in figure 8.

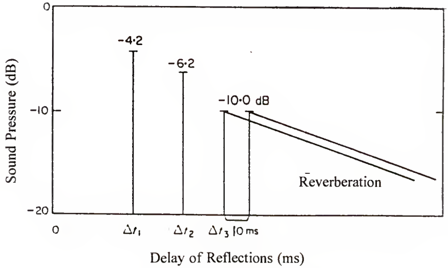



Figure 8 Level and time pattern of reflections and reverberation (Ando, 1979b)

Results. Figure 9 shows the preference scores for varying delay times for each of the four music motifs (*a, b, c,* and *d*). In all but two instances, speaker system *a* (interaural cross correlation = 0.27 to 0.40) received much higher preference scores than speaker system *b* (interaural cross correlation = 0.55 to 0.59).

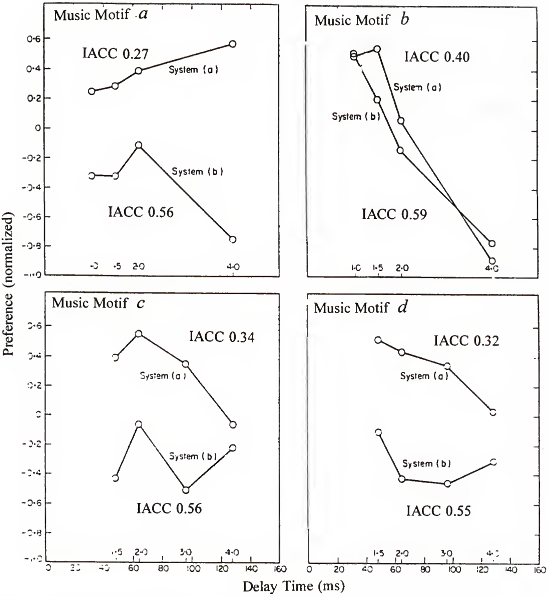



Figure 9 Preference related to reflection delay and interaural cross correlation (Ando, 1979b)

Conclusions. Ando concluded the following at the end of experiment 4. The first experiment showed a large negative correlation between interaural cross correlation and preference scores. Experiment 4 compared spatially different sound fields for different reflection sequences and music motifs. Generally, sound fields with smaller interaural cross correlation values were preferred by listeners.

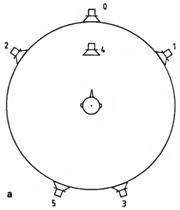
Later Experiments

Ando's later experiments investigated the premise that all of the significant quantitative parameters used to describe the sound at the two ears of a listener in a concert hall can be reduced to four independent factors: level, initial time delay gap, reverberation time, and interaural cross correlation. It was these four parameters that eventually formed Ando's model for predicting listener preference in concert halls.

Method. The method for these studies was nearly identical to those of the previous studies. Sound fields of a concert hall were simulated by computer and presented to multiple listeners in an anechoic chamber using loudspeakers so that the subjects could perform paired comparison tests. Each study fixed two of four parameters at certain values while two other parameters were varied. Within each study, each parameter was evaluated according to its relationship with listener preference and its independence from the other parameters. First, initial time delay gap and reverberation time were varied while level and interaural cross correlation were fixed at constant values. The results of this phase of Ando's study is less relevant to this research project than the succeeding phases. Next, interaural cross correlation and level were varied while initial time delay gap and reverberation time remained fixed. Lastly, Ando varied interaural cross correlation and reverberation time while keeping initial time delay gap and level constant.

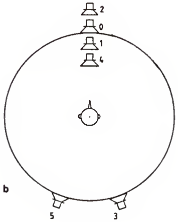
When interaural cross correlation and level were varied, reverberation time was fixed at 3.0 seconds for music motif *A* and 1.0 second for music motif *B*. The initial time delay gap between the direct sound and the early architectural reflections was fixed at 80ms for music motif *A* and 20ms for music motif *B*. Interaural cross correlation was varied by using different loudspeaker configurations around the listeners. Refer to table 1 in the *results* section for the specific interaural cross correlation values for each of the speaker configurations and

music motifs. Figure 10 shows the three different loudspeaker configurations with reflection arrival directions labeled. Figure 11 shows the arrival time and amplitude of the simulated reflections. Four different listening levels were used: 74, 77, 80, and 83dBA.



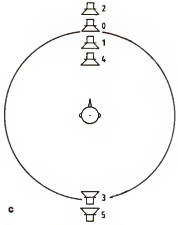

Arrival Direction (degrees)

| | Azimuth | Elevation |
|---|---|---|
| 0 | 0 | 0 |
| 1 | +55 | 0 |
| 2 | -55 | 0 |
| 3 | +160 | 0 |
| 4 | 0 | +12 |
| 5 | -160 | 0 |




Arrival Direction (degrees)

| | Azimuth | Elevation |
|---|---|---|
| 0 | 0 | 0 |
| 1 | 0 | +6 |
| 2 | 0 | -6 |
| 3 | +160 | 0 |
| 4 | 0 | +12 |
| 5 | -160 | 0 |




Arrival Direction (degrees)

| | Azimuth | Elevation |
|---|---|---|
| 0 | 0 | 0 |
| 1 | 0 | +6 |
| 2 | 0 | -6 |
| 3 | 180 | +3 |
| 4 | 0 | +12 |
| 5 | 180 | -3 |

Figure 10 Speaker configurations *a*, *b*, and *c* (Ando,1985)

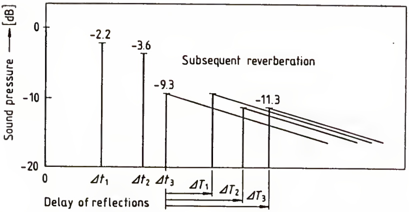



Figure 11 Level and time pattern of reflections and reverberation (Ando, 1985)

Results. Table 1 shows the results of Ando's experiment. Preference scores are given for the various speaker systems, interaural cross correlation values, and listening levels.

Table 1 Preference scores for various speaker systems and listening levels

a) Music motif A

| System | IACC | Listening level [dBA] 74 | 77 | 80 | 83 |
|---|---|---|---|---|---|
| (a) | 0.39 | 0.53 | 0.85 | 0.73 | −0.07 |
| (b) | 0.72 | 0.07 | 0.35 | 0.36 | −0.35 |
| (c) | 0.98 | −0.64 | −0.35 | −0.50 | −0.98 |

b) Music motif B

| System | IACC | Listening level [dBA] 74 | 77 | 80 | 83 |
|---|---|---|---|---|---|
| (a) | 0.42 | 0.10 | 0.55 | 0.92 | 0.20 |
| (b) | 0.67 | −0.35 | 0.27 | 0.37 | 0.04 |
| (c) | 0.98 | −0.85 | −0.38 | −0.41 | −0.46 |

Conclusions. Ando concluded that for a constant listening level, the sound fields with a smaller interaural cross correlation value were always preferred. He also concluded that the preferred listening levels were found in the ranges from 77dBA to 79dBA for music motif *A* and 79dBA to 80dBA for music motif *B*. In addition, results of the analysis of variance clearly showed that both level and interaural cross correlation were independent influences on the subjective preference judgments. Lastly, Ando concluded that the contributions of interaural cross correlation to the changes in preference were larger than those of the listening levels in the ranges tested.

When reverberation time and interaural cross correlation were varied and initial time delay gap and level were fixed, Ando reached similar conclusions. First, the analysis of variance confirmed that both reverberation time and interaural cross correlation independently influenced subjective preference. Second, the preferred listening condition was generally obtained by minimizing the interaural cross correlation value.

Preference Model

Ando's cumulative work resulted in a short list of optimum objectives to be used in the design of concert halls. Ando also related these optimum objectives to their subjectively preferred sound qualities. First, listening level is the primary criterion for listening to sound fields in concert halls. Although preferred levels vary slightly with music motif, the ideal range is centered on 79dBA. Second, reflections should arrive soon after the direct sound. The preferred delay between the direct sound and the early reflections (ITDG) varies according to the autocorrelation of the source signal ($C_A$) and the amplitude of the first reflection (A) such that ITDG = $(1-\log_{10}A)(C_A)$.

Ando's third design objective is the subsequent reverberation time after the early reflections. Preferred values again vary with the autocorrelation of the source signal, but a general guide for orchestra music played in concert halls is a preferred reverberation time between 1.0 second and 2.0 seconds. The fourth and final design objective is incoherence at both ears (the only binaural criterion in Ando's list) which is indicated by low interaural cross correlation values. Ando states that all available data indicates a negative correlation between interaural cross correlation and subjective preference. Ando does not give a preferred value of interaural cross correlation. He simply states that in all of the studies, smaller values of interaural cross correlation were preferred.

## Recent Studies

Beranek (1996) has completed the most recent study relating lower interaural cross correlation values with listening preference. He collected data in thirty-five concert halls. For each receiver position, the 500Hz, 1kHz, and 2kHz octave band interaural cross correlation values were averaged. Then, hall average interaural cross correlation values were calculated by averaging all of the independent receiver positions within each hall. The hall average interaural cross correlation values were then compared to the subjective ratings scored by people listening to actual music performances in the halls.

A general relationship between hall average interaural cross correlation values and the subjective ratings of the concert halls was found. The mean interaural cross correlation value for halls that were subjectively rated *excellent* to *superior* was 0.34. The mean value for halls that were rated *good* to *excellent* was 0.44, and the mean value for halls that were subjectively rated *fair* to *good* was 0.58. Beranek concluded that since interaural cross correlation related so strongly to the subjective ratings of the concert halls, it qualifies as one of the most significant physical attributes for judging the acoustical quality of occupied concert halls.

In addition to the studies completed by Schroeder, Gottlob, Siebrasse, Ando, and Beranek, other studies have found strong relationships between interaural cross correlation and various qualitative indices used by listeners while establishing preference (Morimoto & Iida, 1991 - auditory source width; Hidaka *et al.*, 1991 - spatial impression & diffuseness/envelopment; Ando, 1994 - speech clarity and articulation). However, the published work from the past twenty years lacks studies that relate quantitative values of interaural cross correlation with the architectural features of rooms.

Chiang (1994) recently completed one of the most extensive and detailed studies relating over 65 different architectural features and dimensional relationships of receiver positions

within rooms to quantitative acoustic measures including interaural cross correlation. Data was collected at 81 different receiver locations in 18 different rooms. Individually, most of the parameters accounted for less than 12% of interaural cross correlation variability, and no parameter accounted for more than 25% of the variability. None of the statistical models with multiple parameters accounted for more than 50% of interaural cross correlation variability. In comparison, Chiang accounted for 85% of reverberation time variability. The unexplained variability of interaural cross correlation in that study indicates our lack of knowledge about the relationship between interaural cross correlation and design of architectural enclosures.

The literature shows that lower values of interaural cross correlation are preferred by listeners. However, methods of achieving lower interaural cross correlation values in actual spaces have yet to be defined. Eventually, interaural cross correlation may produce information about qualitative assessment of room acoustics that is as useful to designers as reverberation time. However, the complex relationship between interaural cross correlation and architectural design requires further study.

## MEASUREMENT METHOD

### Introduction

This chapter describes the measurement of real-room binaural impulse responses (including the sound source, binaural receiver, and instrumentation) and the calculation of interaural cross correlation (including filtering and processing). Originally, it was believed that the design of the pinnae and auditory canals in the binaural receiver significantly affected the interaural cross correlation values and could possibly be a source of interaural cross correlation variability. Therefore, emphasis was placed on the design of the binaural receiver and its representation of a human listener. Several experiments establish the level of detail required in the binaural receiver for the measurement of interaural cross correlation.

### Sound Source

Since the available filtering and processing software required an impulsive source, a modified 0.38 caliber revolver firing Remington blanks was used during real-room measurements. The barrel of a standard weapon was removed, and the resulting hole was filled with a conical-shaped plug (Bradley, 1986). The intent was to decrease the directionality of the gun as much as possible.

Using a pistol has both advantages and disadvantages. It is small, inexpensive, mobile, and perhaps the only source capable of producing consistently loud omni-directional impulses. Other impulsive sources such as hand claps, balloon bursts, and electronic impulses presented through loudspeakers are often incapable of consistently emitting omni-directional impulses

that are loud enough (and undistorted) in large concert halls. However, both the frequency response and repeatability of the gun are less than ideal.

Frequency Response

The frequency response curve of a typical Remington blank fired with the revolver is shown in figure 12. The energy is greatest in the 1kHz octave band and less intense in the lower frequencies.

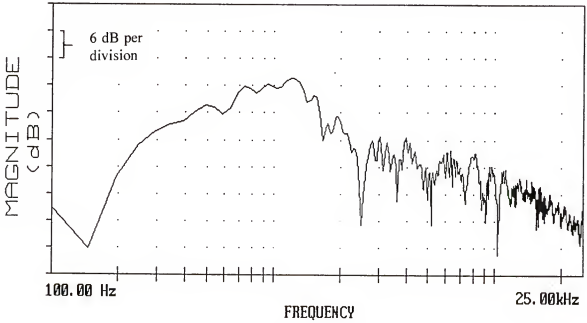



Figure 12 Revolver frequency response curve

The lack of low frequency energy in a gun shot is not a primary concern. In a later section (Interaural Cross Correlation: Sources of Variability in Concert Halls), it will be shown that the variability of interaural cross correlation in the 125Hz and 250Hz octave bands is insignificant compared to the variability at other frequencies. Therefore, the calculation of interaural cross correlation in these octave bands (and consequently the presence of low frequency sound in an impulsive source) is not necessary for this study.

Repeatability

Unfortunately, gunshots are not as consistent in level, directionality, and frequency content as electronic types of sources. As a result, a significant shot-to-shot variability of interaural cross correlation values occurs. Averaging multiple trials for every source/receiver combination can help to decrease the variability due to the measurement system. However, the number of trials is often limited by the time permitted to take all of the acoustical measurements and the number of receiver positions being measured.

Typically, three of four gunshots were fired for each source/receiver combination. Subsequent data analysis showed that even with five shots being averaged, the system variability remained significant. In addition, corrupt or deleted files occasionally resulted in some positions having only two interaural cross correlation values to average. In these cases, the values resulting from any one particular shot had great influence on the average.

Since the variability of the measurement system was significant, it needed to be defined, and then considered when judgements were made about the data. The level of detail permitted during the real room data analysis was consequently limited by the variability due to the measurement system. In other words, slight differences in interaural cross correlation values could have been due to the comparative differences being investigated or the variability of the measurement system. Differences between interaural cross correlation values that were greater than the system variability could be attributed primarily to the characteristics being compared in the study.

Figure 13 shows a representative average of interaural cross correlation values calculated from the impulse responses of four gunshots. Also shown in figure 13 are the confidence intervals which show where 95% of repeated trials would be expected to fall. The width of these confidence intervals indicates the permissible level of detail during data analysis. The

width of the confidence intervals in the lower octave bands (125Hz, 250Hz, and 500Hz) is approximately 0.08 to 0.13, and depends on which four shots are averaged. The width of the confidence intervals in the higher octave bands (1kHz, 2kHz, and 4kHz) is approximately 0.05 to 0.10, and also depends on which four shots are averaged.

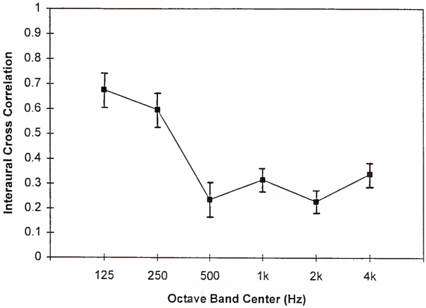



Figure 13 System variability measured using 95% confidence intervals

During data analysis, if interaural cross correlation values differed by amounts less than the width of the corresponding confidence interval, results were judged somewhat inconclusive. It seemed appropriate to state that the values were similar, but they could not be labeled equal. They may be equal, but equality could not be determined considering the variability due to the measurement system.

Unfortunately, five shots was the maximum number of trials at any one position. Therefore, the actual number of shots needed to eliminate the measurement system as a source of variability could not be established. If a gun is to be used in any future research, preliminary studies should investigate the number of shots needed to eliminate the system as a source of variability.

## Binaural Receiver

### Construction

A common department store manikin, designed to represent an adult male, was obtained. The hollow fiberglass torso was filled with an expanding insulation to keep it from resonating at certain frequencies. The fiberglass neck was replaced with a photographic tripod knuckle, allowing the head of the manikin to turn 360 degrees in azimuth, and from -60 degrees to +90 degrees in elevation. The head can also be tilted left to -90 degrees or right to +90 degrees. The crown of the head was made into a gasketed removable plate, allowing access to the instrumentation. The pinnae for the manikin are made of silicon rubber molded from real human ears (judged average in shape and size by a consulting audiologist - Scroggie, 1992) (refer to figure 14).

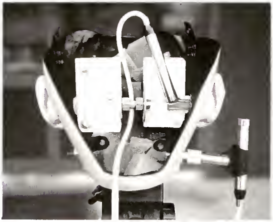


Figure 14 Full scale manikin design and instrumentation detail

The inner ears of the manikin are comprised of three elements; microphones (1/2"Bruel & Kjaer), auditory canals, and housings that couple the microphones to the canals. The ear canals of the manikin are made of vinyl tubing and approximate human ear canals by having a 0.75cm inside diameter, 2.20cm length, and 1.00cc volume (Baur, 1967). The canals extend through the fiberglass shell of the head far enough to contact the pinnae which are glued to the

exterior of the shell. Within the housings, the canals are angled up five degrees from the ear canal openings, and the microphones are angled seventy degrees from horizontal. The canal angle simulates that found in a human ear, and the microphone angle places the transducer in a position similar to that of the tympanic membrane. The two housings are separated by an expansion apparatus that resiliently mounts the instrumentation inside the head. Only the small piece of vinyl canal that extends through the shell to the pinna actually comes in contact with the shell.

Testing

The head related transfer functions of the manikin were measured at the University of Florida's Departments of Neuroscience and Surgery by Dr. John C. Middlebrooks who had previously performed the same measurements on multiple human subjects (Middlebrooks *et al.*, 1989 and Middlebrooks and Green, 1990). The subjects were seated in an anechoic chamber. The stimuli were presented using dynamic loudspeakers (Radio Shack cat. No. 40 1289A) which were positioned on a 1.20m semicircular hoop frame at ten degree intervals. The frame was mounted vertically with its diameter perpendicular to and passing through the center of the subject's interaural axis. Horizontal source positions were varied by moving the frame (using a computer-controlled stepper motor) about its vertical axis in 10 degree increments. Vertical source positions were varied by activating different loudspeakers along the frame using a computer-controlled multiplexor.

The sound source positions are defined in a double-pole coordinate system (Knudsen, 1982). The horizontal position is noted as the azimuth angle so that 0 degrees is forward, +90 degrees is to the side of the microphone, and -90 degrees is to the opposite side of the microphone. Azimuth positions ranged from -160 degrees to +160 degrees. The vertical location is noted as the elevation angle and is measured relative to a horizontal plane.

Positions below the horizontal plane containing the subject's interaural axis are denoted with negative values while those above are denoted with positive values. Elevation positions for this experiment ranged from -40 degrees to +90 degrees (refer to figure 15). Several example source positions with their corresponding coordinates are listed below.

| Azim. | Elev. | Position |
|---|---|---|
| 0 | 0 | 1 |
| +90 | 0 | 2 |
| +90 | 60 | 3 |

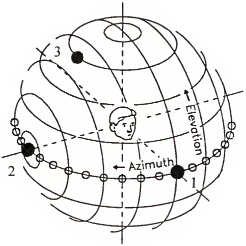



Figure 15 Spatial coordinate system of source (Middlebrooks *et al.,* 1989)

The stimulus waveform was produced using an inverse fast Fourier transform. The duration from onset to offset was 10.20ms The stimulus bandwidth was 1kHz-14kHz with components spaced 97.7Hz apart. Frequency components outside the stimulus bandwidth were more than 30dB down. At each loudspeaker position, the stimulus was repeated 16 times with 10.24ms gaps separating them. The 16 measured responses were then averaged.

When human subjects were used, sounds were recorded with a Knowles model EA-1934 microphone placed at least 5.0mm into the auditory canal. The manikin had a 1/2" Bruel and Kjaer microphone inside the right ear housing (truncating the 2.20cm ear canal). The output signals were then filtered between 1kHz and 14kHz, amplified, and digitized at a sampling rate

of 50kHz with 12 bit resolution. A total of 324 source positions were measured for each subject. Each data set took only six minutes to measure. The subjects were instructed to refrain from moving or swallowing during the measurement session. If the subject's head moved more than 2 degrees in any direction (as indicated by an electromagnetic sensor) the data were rejected. (Refer to Middlebrooks *et al.*, 1989 for a more detailed description of measurement methods.)

## Results

The manikin's head related transfer function (azimuth 0, elevation 0) was compared to that of the Bruel and Kjaer Head & Torso Simulator (type 4128) (refer to figure 16). In addition, the manikin data were compared to the human subject data previously measured by Middlebrooks (refer to figure 17). The head related transfer functions (azimuth +90, elevation 0) in figure 17 do not contain ear canal resonances. They were removed from the data sets so that inconsistent microphone positioning within the auditory canals of the manikin and human subjects was not a source of variability. (A detailed description regarding the removal of the ear canal resonances can be found in Middlebrooks and Green, 1990.)

## Intersubject Variability

In most instances, the manikin's data were within 5.0dB of the subject's data. However, there were places where the variability was greater. Middlebrooks confirmed that the same degree of variability (*i.e.,* the difference in the sound pressure levels in dB) are common among human subjects, especially at higher frequencies (1kHz and above). Figure 18 shows the large amount of intersubject variability of ten human subject head related transfer functions (azimuth +0, elevation 0).

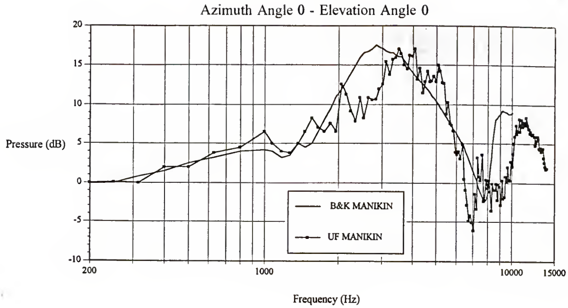



Figure 16 Manikin / manikin head related transfer function comparison

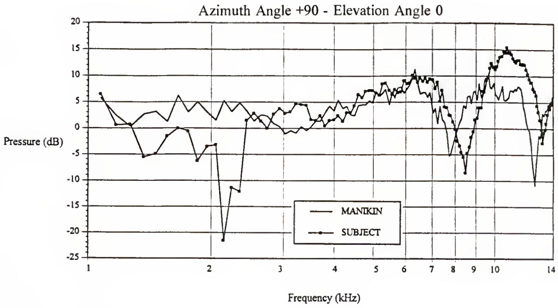



Figure 17 Manikin / human subject head related transfer function comparison

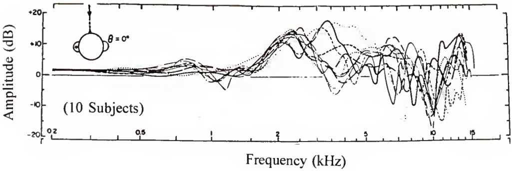



Figure 18 Intersubject variability between ten human head related transfer functions (Shaw, 1966)

The most common type of variability in the data occurred when the plotted curves for the manikin and subject were similar in shape (amplitude), but were shifted laterally relative to each other (refer to figure 19). This displacement caused corresponding maxima and minima to occur at different frequencies. This type of variability also occurs between human subjects (Middlebrooks *et al.*, 1989).

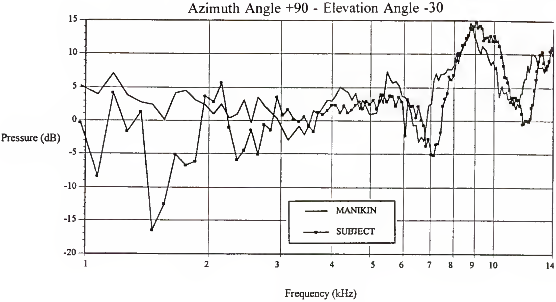



Figure 19 Variability between manikin and human subject head related transfer functions

Middlebrooks explained that response patterns previously measured for multiple human subjects were all qualitatively similar to each other, however they occurred at different frequencies for different subjects. For example, subject 1 had a specific response pattern at 8kHz recognizable by two discrete amplification maxima. Both maxima were centered at an azimuth angle of +90 degrees, but one occurred above the interaural horizontal plane while the other occurred below (refer to figure 20c - The shaded spheres in figure 20 represent human subject directional sensitivity. Maxima regions, where hearing is most sensitive, are shaded. Contours represent attenuation relative to maximum measured values). Subjects 2 and 5 predictably had a similar response pattern, but occurring at different frequencies (6kHz and 8.9kHz respectively) (refer to figures 20a & 20d). When the response patterns for the same subjects were compared at one consistent frequency, the intersubject variability produced quite different response patterns (refer to figures 20c & 20b, also 20c and 20e).

Middlebrooks explained that there is a general correlation between the frequencies at which particular patterns occur and the physical sizes of the subjects. Smaller subjects tend to have patterns occurring at higher frequencies than larger subjects. In the previously conducted work, subject 1 (pattern at 8kHz) was 183cm tall, while subject 2 (similar pattern at 6.9kHz) was 193cm tall. Subject height was used only as a general judge of overall subject size.

Therefore, part of the variability in the manikin versus subject head related transfer function comparisons can be explained by varying physical sizes. Smaller differences (less than 5dB and unrelated to size differences) between the manikin and human subject head related transfer functions can be explained by the subtle differences of facial detail and pinnae shape. However, there were specific instances where the subject's head related transfer functions decreased drastically in level (for example, -17db at 1.5kHz. in figure 19). Middlebrooks stated that decreases such as these are unusual at these frequencies (refer to the

ten other human head related transfer functions in figure 18) and were not explainable. Middlebrooks did confirm that the manikin's head related transfer functions should not necessarily decrease like the subject's head related transfer functions in these particular instances.

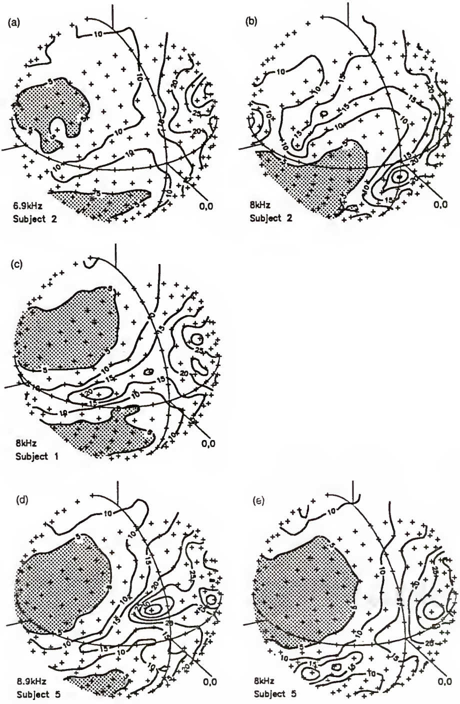



Figure 20 Human subject directional sensitivity plots (Middlebrooks *et al.*, 1989)

Since the manikin's head related transfer functions were similar to those of a human subject and to those of a commercially available head and torso simulator, it was concluded that the manikin was representing an actual human head will sufficient accuracy, especially since the variability between multiple human head related transfer functions often exceeds that which occurred between the manikin and human subject used in this research.

## Instrumentation, Filtering, and Processing

The manikin is equipped with two 1/2" Bruel & Kjaer microphones and pre-amplifiers. The signals are fed directly into a multi-channel 12 bit LeCroy digitizer (model 6810) where they are sampled at a rate of 50kHz. Using Catalyst Waveform software (commercially marketed with the LeCroy digitizer), the two binaural signals are acquired and then written to computer storage for later processing.

Octave band digital filtering and calculation of interaural cross correlation is done using the ARIAS (Acoustical Research Instrumentation for Architectural Spaces) software package previously developed at the University of Florida by Dr. Harold Doddington and Dr. Bill Schwab. ARIAS uses equation 5 to calculate interaural cross correlation.

Equation 5

$$IACC_{(\max)} = \frac{\int_{t_1}^{t_2} P_{lt}(t) P_{rt}(t+\tau)\,dt}{\left(\int_{t_1}^{t_2} P_{lt}^2\,dt \int_{t_1}^{t_2} P_{rt}^2\,dt\right)^{1/2}}$$

$IACC$ Maximum interaural cross correlation value.
$P_{lt}$ Sound pressure of the left binaural signal at time *(t)*.
$P_{rt}$ Sound pressure of the right binaural signal at time (*t*).
*tau* Amount of time that the right signal is shifted relative to the left signal before IACC is calculated. A range of values from -1.0ms to +1.0ms at 20 microsecond intervals is used, totaling 100 calculated IACC values. Then, the maximum absolute value is used.
$t_1$ This value is 0.0ms, the instant when the direct sound reaches the triggering microphone.
$t_2$ This value is set at 80ms.

## Measurement Method Test

To test both the manikin design and the measurement method, binaural impulse responses were collected in conjunction with the National Research Council of Canada in multiple concert halls at designated receiver positions. The Canadian research team, headed by Dr. John Bradley, used a commercially marketed Bruel & Kjaer Head & Torso Simulator (type 4128) and the RAMsoft software system. The sound source was an MLS signal played through a dodecahedron loudspeaker. The resulting interaural cross correlation values from the Canadian research team were then compared to those measured using the modified pistol, the constructed manikin, and the ARIAS software system.

Figure 21 shows interaural cross correlation values measured using the two different systems for a main floor seat inside the Philadelphia Academy of Music. The data collected by the University of Florida research team is shown with the 95% confidence intervals that indicate the measurement system variability. The measurement system of the Canadian research team has an insignificant amount of trial-to-trial variation, so confidence intervals are not necessary. Since the data measured by the Canadian research team typically fell within the confidence intervals of the data measured by the University of Florida research team, it can be concluded that the interaural cross correlation values measured by the two teams were generally similar.

Statistical correlations of interaural cross correlation values measured by the two teams resulted in the following correlation coefficients (r-values at 5% significance): 500Hz octave band +0.66, 1kHz octave band +0.86, 2kHz octave band +0.75, and 4kHz octave band +0.70. One would expect that two quantitative measurement systems would relate more consistently than the ones in this study. Reasons for the interaural cross correlation differences still require further investigation. The use of a gunshot as an impulsive source by the University of

Florida research team is a probable cause of the differences between the values measured by the two systems. As previously explained, the gunshot varies from shot to shot. Typically only three or four shots were averaged for each receiver position. If for any reason one of the shots varied significantly from the others at any single position, the average was still greatly affected. Subsequent measurement method studies should perhaps begin with using a more similar and consistent sound source. The following two experiments show that the design of the manikin and the location of the microphones within the manikin were not causes of the differences between the interaural cross correlation values measured by the two research teams.

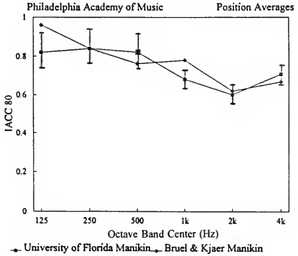



Figure 21 Interaural cross correlation measurement method comparison

## Microphone Location Experiment

This experiment investigated whether octave band (125Hz - 4kHz) interaural cross correlation values measured with the microphones inside the manikin's head (truncating the 2.2cm long auditory canals) were significantly different from those taken with the microphones outside the head approximately 30mm from the ear canal openings. The experiment was originally intended to discover if data taken prior to the manikin's construction (with a seated person holding two omni-directional microphones next to the left and right ears) was

comparable to that taken with the microphones placed inside the head of a detailed manikin. In addition, the effect of the filtering characteristics of the pinnae and auditory canals on interaural cross correlation was studied.

For this experiment, the manikin was equipped with four microphones (all 1/2" Bruel & Kjaer). Two were inside the head at the left and right ear drum locations, while the other two were mounted outside the manikin's head approximately 30mm from each of the ear canal openings. Binaural impulse responses, recorded with the microphones positioned inside and outside the manikin head, were measured at various receiver positions in multiple concert halls. Three different methods were used to compare the signals from the two microphone locations. First, interaural cross correlation values were calculated and statistically compared. Second, the right ear signals recorded inside the head were cross correlated with right ear signals recorded outside the head. Similarly, left ear signals recorded inside the head were cross correlated with left ear signals recorded outside the head. Lastly, the actual signals (pressure as a function of time) for different frequencies (500Hz and 4kHz) were plotted and visually compared.

Statistical comparisons of means were performed for different receiver positions, source positions, concert halls, and integral durations ($t_2$ = 30ms, 50ms, 80ms, 100ms). Ninety-eight percent of the comparisons showed that there were no statistical differences between octave band interaural cross correlation values taken at the inside and outside microphone positions. Figure 22 compares interaural cross correlation values from Troy Music Hall measured with the microphones located inside the manikin head against those measured with the microphones located outside the manikin head. The confidence intervals indicating system variability are not shown since the slight differences between microphone locations are less than the confidence intervals established in figure 13.

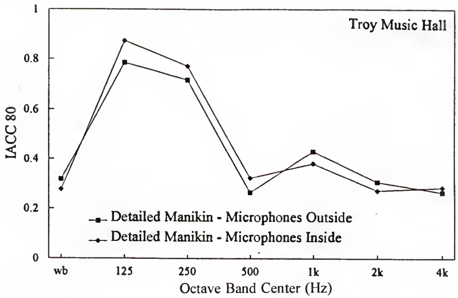
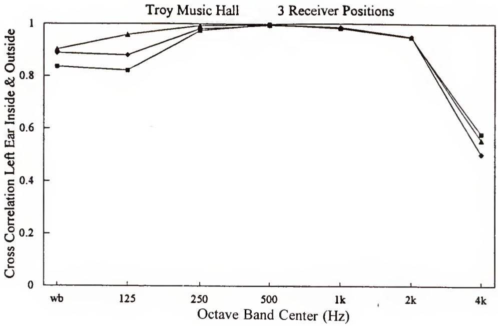



Figure 22 Microphone location comparison

When signals recorded inside and outside the same ear were cross correlated, results showed that the signals were similar in the 125Hz octave band (cross correlation values around 0.85-0.95), identical in mid frequency octave bands (250Hz, 500Hz, 1kHz, 2kHz) (cross correlation values around 0.95-1.0), and dissimilar in the 4kHz octave band (cross correlation values around 0.5) (refer to figure 23).

It is believed that the filtering effects of the pinnae and head as well as the resonances caused by the longitudinal fundamental frequency of the manikin's auditory canals were the causes of the dissimilarity between the 4kHz octave band signals recorded at the two microphone positions. Visual inspection of the 500Hz octave band signals (plotted as pressure versus time) reconfirms that the signals recorded inside the head are similar to the those recorded outside the head. Visual inspection of the 4kHz signals shows that signals recorded inside the head are dissimilar to those recorded outside the head (refer to figures 24 & 25). It is interesting to note that even when the 4kHz octave band signals recorded inside and outside

## Left Ear




## Right Ear

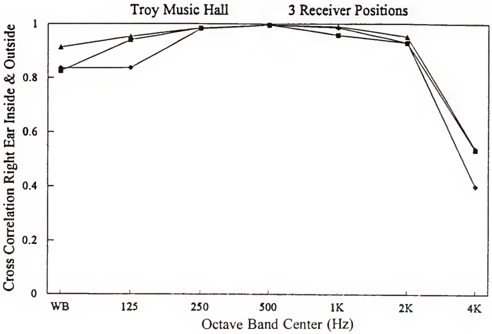



Figure 23 Correlation of signals measured inside and outside the manikin head

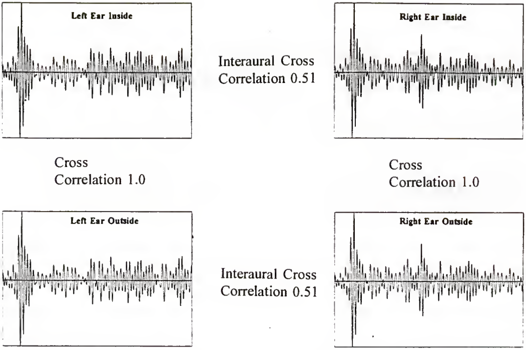



Figure 24 Signal comparison - 500Hz octave band

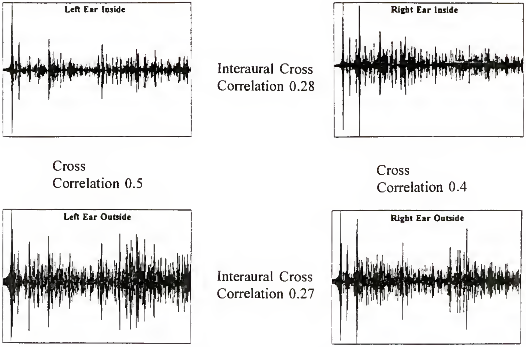



Figure 25 Signal comparison - 4kHz octave band

the same ear differ by as much as a 0.4 cross correlation value, the interaural cross correlation values are not significantly different (0.28 when the microphones were located inside the head and 0.27 when the microphones were located outside the head).

It was concluded that despite the evident high frequency effects of the pinnae and auditory canals, microphone positioning inside or outside the head does not affect octave band interaural cross correlation values. Therefore, future measurements of interaural cross correlation do not necessitate a manikin with detailed pinnae and auditory canals.

## Level of Detail Experiment

This experiment investigated whether the general head size, amount of facial detail, and presence of detailed pinnae and auditory canals on a manikin affect octave band interaural cross correlation values. Binaural impulse responses were recorded in the Center for the Performing Arts in Gainesville, Florida. This room is a medium-sized, multi-use music hall with one large balcony. It seats approximately 1800 people, and its mid-frequency unoccupied reverberation time is 1.8 seconds. Impulse responses were recorded at one receiver position located on the main floor, midway back, and approximately twenty-five feet from the hall's longitudinal axis. The following six receiver configurations were used to record the impulse responses (refer to figure 26 and table 2):

| Config. | Description |
| --- | --- |
| 1 | Two omni-directional microphones (1/2" Bruel &Kjaer) placed 112cm above the floor and spaced 20cm apart. |
| 2 | Same as 1, except a common football was placed between the microphones. |
| 3 | Same as 1, except an abstract manikin head having little facial detail and no pinnae or auditory canals was placed between the microphones. |
| 4 | Same as 1, except a manikin with a lot of facial detail and fairly accurate pinnae (no auditory canals) was placed between the microphones. |
| 5 | Same as 1, except a manikin with an average amount of facial detail and highly detailed pinnae and auditory canals was placed between the microphones. |
| 6 | Same as 5 , except the microphones were placed inside the manikin head truncating the 2.20cm auditory canals. |

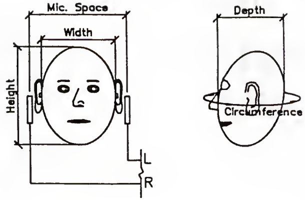



Figure 26 Manikin measurement method

Table 2 Receiver characteristics

| | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| Head Width | - | 16.5cm | 14.5cm | 13.2cm | 14.0cm | 14.0cm |
| Head Depth | - | 17.0cm | 19.0cm | 18.3cm | 20.0cm | 20.0cm |
| Head Height | - | 26.0cm | 21.0cm | 21.5cm | 22.0cm | 22.0cm |
| Circumference | - | 52.5cm | 55.2cm | 54.0cm | 59.0cm | 59.0cm |
| Mic. Spacing | 20.0cm | 20.0cm | 20.0cm | 20.0cm | 20.0cm | - |
| Mic. to Head Dist. | - | 1.75cm | 2.75cm | 3.40cm | 3.00cm | - |
| Facial Detail | - | None | Low | High | Medium | Medium |
| Pinnae Detail | - | None | Low | Medium | High | High |
| Auditory Canals | - | No | No | No | Yes | Yes |

Octave band interaural cross correlation values were calculated for each receiver configuration. Results were not as conclusive as those for the microphone location experiment. Statistical comparison of means (5% significance level) primarily showed that receiver configuration *1* (no manikin placed between the two microphones) resulted in values that were significantly different from values measured using the other receiver configurations (refer to figure 27). However, isolated values measured with some of the receiver configurations were significantly different than certain values measured with other receiver configurations, but only at some frequencies. It is not known if the differences between the

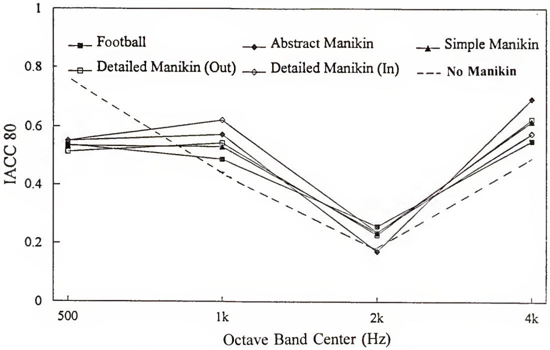



Figure 27 Effect of receiver configuration on interaural cross correlation

values are due to the variability of the measurement system or the level of detail in the receiver. Given the uncertainty of the sources of the variability between the measured values, it can only be said that the various receiver configurations resulted in similar interaural cross correlation values, and the experiment should perhaps be redone using a measurement system with less variability.

## INTERAURAL CROSS CORRELATION: SOURCES OF VARIABILITY IN CONCERT HALLS

### Data Collection

During the summer of 1992, The Concert Hall Research Group (an organization of professional acoustical consultants, researchers, and scholars) sponsored a trip to take acoustical measurements in multiple Northeastern United States concert halls by three independent research teams. The primary purpose was to standardize acoustical measurement methods while establishing a database for future research. Two of the measurement teams, one from the University of Florida and the other from the National Research Council of Canada, measured binaural impulse responses and calculated interaural cross correlation at 84 independent receiver locations. A list of the measured halls appears in table 3. Floor plans and sections of the halls which show sound source and receiver positions appear in appendix *A*. The measurement methods of the two research teams have already been described in the *Measurement Method* section (pg. 24).

Table 3 Measured concert halls

| Name | Location | Room Volume (cf) | No. Seats | Reverberation Time (sec.) (mid-frequency) |
|---|---|---|---|---|
| Boston Symphony Hall | Boston, MA | 670,579 | 2,555 | 2.40 |
| J.F. Kennedy Center | Washington, D.C. | 763,501 | 2,759 | 1.66 |
| Kleinhans Music Hall | Buffalo, NY | 644,000 | 2,839 | 1.58 |
| Meyerhoff Concert Hall | Baltimore, MD | 756,558 | 2,465 | 2.05 |
| Orchestra Hall | Detroit, MI | 577,676 | 2,038 | 1.65 |
| Philadelphia Academy of Music | Philadelphia, PA | 554,861 | 2,914 | 1.17 |
| Severance Hall | Cleveland, OH | 554,000 | 1,996 | 1.55 |
| Troy Savings Bank Music Hall | Troy, NY | 399,778 | 1,097 | 2.31 |

Preliminary Data Analysis

The interaural cross correlation values from all receiver positions were averaged, and the standard deviation of values at all positions was calculated (standard deviation is used as a measure of variability about the average values) (refer to figures 28 and 29).

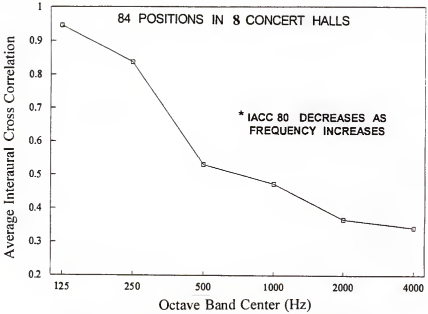



Figure 28 Average interaural cross correlation values

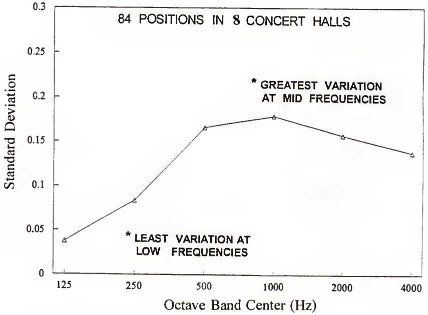



Figure 29 Standard deviation of interaural cross correlation values

Figure 28 shows that on average, interaural cross correlation decreases as frequency increases. Figure 29 shows that the greatest amount of variability occurs in the 500Hz, 1kHz, and 2kHz octave bands. Variability of interaural cross correlation values in the 125Hz and 250Hz octave bands is minimal, and was judged to be insignificant relative to the other octave bands.

The hall average for each of the eight halls was calculated by averaging the independent positions within each hall. The standard deviation of all positions within each hall was calculated and used as a measure of within-room variability. Lastly, the standard deviation of independent positions common to all halls was calculated, and used as a measure of among-room variability. Among-room variability of hall average interaural cross correlation was relatively small. This means that the difference between the highest hall average and lowest hall average was not large. On average, among-room variability of positions common to all halls is less than the average within-room variability of independent positions (refer to figure 30). For example, the amount of within-room variability inside Boston Symphony Hall exceeds the amount of variability for a common position in all nine halls (regardless of the position). This means that there are generally greater differences between the multiple positions within one room than there is between the same position in all of the eight concert halls.

As shown in figure 30, within-room variability of interaural cross correlation exceeds the other two types of variability. For this reason, within-room variability served as the starting point for further investigation. Data analysis showed that for seats within one room, those close to the stage typically had high interaural cross correlation values, while those further from the stage typically had low interaural cross correlation values (refer to figure 31).

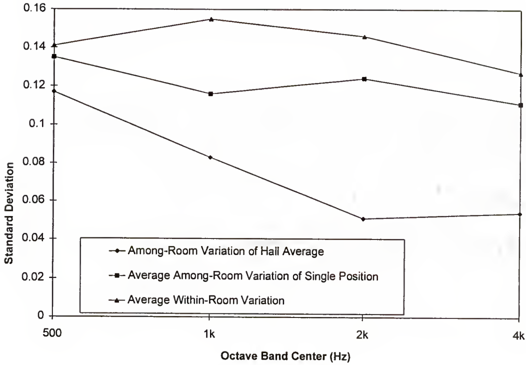



Figure 30 Comparison of standard deviations measured using different sets of interaural cross correlation values

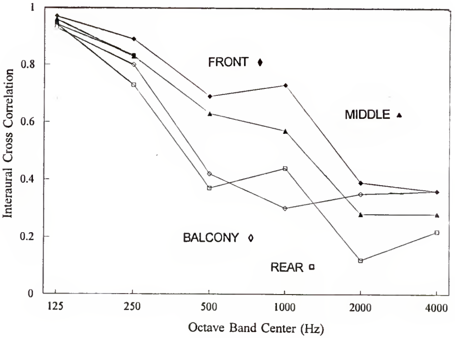



Figure 31 Decrease of interaural cross correlation values with distance from the source

It was suspected that there might be a relationship between the distance of the receiver position from the source and interaural cross correlation values. However, seats in the front of upper balconies, despite being further from the stage, often had interaural cross correlation values higher than those measured at seats in the rear of the main floor. This discrepancy indicated that the general decrease in interaural cross correlation as one moves back in a concert hall was probably not due to the simple linear distance from the sound source.

Plotted impulse responses and the corresponding interaural cross correlation values for certain balcony positions were compared to those for main floor seats that were slightly closer to the stage but still in the rear of the room. The comparison indicated that the higher interaural cross correlation values in the balcony seats was a result of the level of the direct sound relative to the level of the architectural reflections. The balcony seats typically had a larger ratio of direct sound to early reflected sound than the seats in the rear of the main floor (as long as the balcony fasciae did not obstruct the direct sound paths to the balcony seats). Since the direct sound was relatively higher in level and reached the two ears simultaneously, it made the left and right signals more similar, and tended to increase the overall interaural cross correlation value.

The direct sound to the seats in the rear of the main floor was lower in level relative to the architectural reflections due to the attenuation from grazing over the absorbent seat backs. Therefore, the direct sound did not contribute as much in making the left and right signals similar. As a result, the overall interaural cross correlation value was lower. It seemed then, that the ratio of the direct sound to the early reflected sound was naturally higher for balcony seats than it was for seats in the rear of the main floor. It was suspected that this ratio between the direct and early reflected sound energy may relate more strongly to interaural cross correlation values than the simple linear distance from the sound source.

## Direct to Reflected Energy Ratio

It was found that the variability of interaural cross correlation within and among rooms is significantly related to the level of the direct sound relative to the level of the architectural reflections. Receiver positions close to the sound source typically have a direct sound that is high in level. These positions also typically have high interaural cross correlation values. After experiencing air absorption, natural geometric spreading, attenuation from passing over seat backs, and diffraction over balcony fasciae, the direct sound for receiver positions in the rear of the room is typically weak. Often, positions in the rear of the balconies or at the rear of the main floor have a direct sound that is lower in level than the multiple succeeding architectural reflections and a low interaural cross correlation value.

Equation 6

$$D/R_{(mon)} = \frac{\int_{t_1}^{t_2} P^2(t)dt}{\int_{t_2}^{t_3} P^2(t)dt}$$

*D/R* Direct to reflected energy ratio of a monaural signal.
*P(t)* Sound pressure of a monaural signal at time *(t)*.
$t_1$ This value is 0.0ms, the instant when the direct sound reaches the triggering microphone.
$t_2$ This value is 5.0ms. The integral duration from $t_1$ to $t_2$ includes the direct sound, but not the architectural reflections.
$t_3$ This value is 80.0ms. The integral duration from $t_2$ to $t_3$ includes the architectural reflections, but not the direct sound.

Since $t_3$ in the direct to reflected energy ratio equation (6) and $t_2$ in the interaural cross correlation equation (5) are both 80ms, a more accurate description of the study would be the effect of the direct to early reflected energy ratio on the variability of early interaural cross correlation. However, for simplicity, the study will refer to early interaural cross correlation where $t_2$ = 80ms as IACC and the direct to early reflected energy ratio where $t_3$ = 80ms as the D/R energy ratio.

Initial Correlation of Interaural Cross Correlation and Direct to Reflected Energy Ratios

An initial test was performed to see if IACC values related to the corresponding D/R energy ratios. The resulting pearson correlation coefficient (r-value) (5% confidence level) was 0.57 for 205 samples (41 positions, 5 frequencies, 7 halls - Meyerhoff Hall was excluded from this part of the study due to questionable data acquisition).

Several preliminary conclusions can be based on this correlation coefficient. First, there is a relationship between IACC and D/R energy ratios. When the correlation coefficient (*r*-value) is squared ($r^2$-value) the resulting amount of explained variability is 33%. In other words, one third of IACC variability within and among rooms is due simply to the level of the direct sound relative to the level of the architectural reflections. To establish the significance of this amount of variability, recall that Chiang (1994) was only able to explain 12% of interaural cross correlation variability with any single architectural feature (Refer to the *Relationship Between Interaural Cross Correlation and Listening Preference* section, pg. 22).

Subgrouping by Frequency

The next step in the data analysis was to establish if the effect of the direct sound on IACC varied with frequency (*i.e.*, the correlation coefficient between IACC and D/R energy ratios varied when the data was subgrouped by octave band). The resulting correlation coefficients are given in table 4.

Table 4 Correlations between interaural cross correlation and direct to reflected energy ratios (subgrouped by frequency)

| | | |
|---|---|---|
| All Data | 0.57 | *205 samples* |
| 500Hz | 0.27 | *41 samples* |
| 1kHz | 0.70 | *41 samples* |
| 2kHz | 0.76 | *41 samples* |
| 4kHz | 0.44 | *41 samples* |
| WB | 0.79 | *41 samples* |

Several important conclusions can be made from the correlation coefficients in table 4. The relationship between IACC and D/R energy ratios does seem to vary with frequency. Data show that the correlation is much stronger in the 1kHz and 2kHz octave bands than in the 500Hz and 4kHz octave bands. In fact, well over half (50%-60%) of the 1kHz and 2kHz IACC variability is due simply to the level of the direct sound relative to the level of the architectural reflections. Further research is suggested to discover why this effect changes with frequency. However, since the changing effect of the direct sound with frequency is essentially a monaural source of variability, it did not merit further investigation for the purposes of this research project. Instead, focus was placed on the sources of variability that produced binaural information about concert hall acoustics.

<u>Initial Subgrouping by Hall</u>

Since the relationship between IACC and D/R energy ratios varied with frequency, the next step of the data analysis was to investigate if the relationship also varied among halls. However, when subgrouping by hall, the strength of the relationship between the D/R energy ratios and the within-room variability of IACC depended greatly upon which frequency subgroups were included in the data set. Table 5 shows how the relationship between the D/R energy ratios and the within-room variability of IACC changes according to which frequency subgroups are included in the data set. Basically, the coefficients increase when the less related frequency bands (*i.e.*, 500Hz and 4kHz octave bands) are removed from the data set.

It is important to note that regardless of which frequency subgroups are included in the data set, the correlation coefficients seem to vary among halls. Also of interest is the fact that the ranking of each individual room relative to the others remains similar regardless of the frequency subgroups. In other words, in each of the columns, Boston Symphony Hall receives the lowest *r*-value while the Kennedy Center receives the highest.

Table 5 Correlations between interaural cross correlation and direct to reflected energy ratios (subgrouped by hall with various frequency subsets)

| | 500, 1k, 2k, 4k | WB, 500, 1k, 2k, 4k | WB, 1k, 2k |
|---|---|---|---|
| Boston | 0.26 | 0.33 | 0.42 |
| Kennedy | 0.75 | 0.78 | 0.88 |
| Troy Music | 0.46 | 0.51 | 0.64 |
| Philadelphia | 0.51 | 0.54 | 0.47 |
| Kleinhans | 0.57 | 0.60 | 0.78 |
| Severance | 0.57 | 0.61 | 0.61 |
| Orchestra | 0.55 | 0.62 | 0.70 |

Subgrouping by Frequency and Hall

Since the strength of the relationship between IACC and the D/R energy ratios vary with both frequency and hall, the data should ideally be subgrouped by frequency and by hall. However, if this is done the number of samples decreases drastically. For example, if the 1kHz IACC values in Boston Symphony Hall were to be correlated with the 1kHz D/R energy ratios of the same hall, only six samples could be used. Despite the small sample sizes, table 6 shows how subgrouping the data by frequency and hall could allow studies of greater detail to occur. If additional data were collected, and added to the existing data, perhaps more could be learned about why the effect of the direct sound on interaural cross correlation varies with both frequency and hall.

Table 6 Correlations between interaural cross correlation and direct to reflected energy ratios (subgrouped by hall and frequency)

| | WB | 500Hz | 1kHz | 2kHz | 4kHz |
|---|---|---|---|---|---|
| Boston | 0.86 | -0.65 | 0.40 | 0.73 | -0.11 |
| Kennedy | 0.96 | 0.66 | 0.89 | 0.92. | 0.74 |
| Troy Music | 0.75 | -0.59 | 0.74 | 0.44 | 0.12 |
| Philadelphia | 0.72 | 0.21 | 0.54 | 0.44 | 0.51 |
| Kleinhans | 0.90 | 0.30 | 0.87 | 0.86 | 0.60 |
| Severance | 0.91 | 0.37 | 0.62 | 0.38 | 0.24 |
| Orchestra | 0.96 | -0.42 | 0.63 | 0.67 | 0.81 |
| Average | 0.87 | -0.02 | 0.67 | 0.63 | 0.42 |

Table 6 shows how the effect of the direct sound varies both with frequency and hall. However, since subgrouping the data by frequency and hall decreases the sample size beyond statistical confidence, either more data needs to be collected in each of the rooms or the available data must be grouped differently. Therefore, when studying the effect of the direct sound versus frequency, rooms in the existing sample must be grouped together. It must be understood that any particular room may have an independent relation higher or lower than that of the whole group of rooms. However, to compare the effect of the direct sound for one octave band relative to another, the method must suffice.

Similarly, when studying the effect of the direct sound on the within-room variability of IACC, frequencies must be grouped together. Again, it must be understood that any particular frequency may have an independent relation higher or lower than that of the whole group. For example, the 500Hz, 1kHz, 2kHz, and 4kHz octave band IACC values in Boston Symphony Hall may correlate with their corresponding D/R energy ratios producing a $r$-value of 0.45. However, if more data allowed subgrouping by frequency also, the independent frequency coefficients may be as follows: 500Hz (0.12); 1kHz (0.88); 2kHz (0.93); 4kHz (0.39).

## Subgrouping by Hall

The final step in the analysis was to subgroup the data set by hall using specified frequencies (500Hz, 1kHz, 2kHz, & 4kHz) in order to increase the sample size (refer to table 7). As already explained, if the 500Hz or 4kHz values are excluded from the data set all of the correlation coefficients become higher.

The subgroups in table 7 have been presented in order of decreasing correlation coefficients. The Kennedy Center, Kleinhans Music Hall, Severance Hall, and Detroit's Orchestra Hall are above average while The Philadelphia Academy of Music, Troy Music Hall, and Boston Symphony Hall are below average. It can be concluded that the D/R energy

ratio is related to IACC (refer to figure 32), especially in rooms such as The Kennedy Center and Kleinhans Music Hall and in the 1kHz and 2kHz octave bands. Figure 32 shows that 58% of interaural cross correlation variability in The Kennedy Center is explained by D/R energy ratios.

Table 7 Correlations between interaural cross correlation and direct to reflected energy ratios (subgrouped by hall) (Data Set 500Hz, 1kHz, 2kHz, 4kHz)

| | | |
|---|---|---|
| Kennedy | 0.75 | 28 samples |
| Kleinhans | 0.57 | *24 samples* |
| Severance | 0.57 | *24 samples* |
| Orchestra | 0.55 | *28 samples* |
| *Average* | *0.52* | |
| Philadelphia | 0.51 | *24 samples* |
| Troy | 0.46 | *16 samples* |
| Boston | 0.26 | *20 samples* |

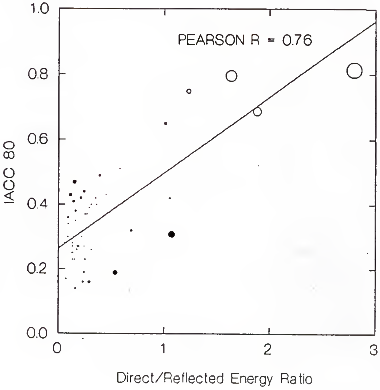



Figure 32 Interaural cross correlation values measured inside the Kennedy Center correlated with direct to reflected energy ratios

## Excluding the Direct Sound

After the relative level of the direct sound was identified as a source of variability that provided no useful binaural information about the architectural reflections, interaural cross correlation values were recalculated using an integral duration that started at $t_1$ = 5ms. It was believed that starting the integration 5ms after the arrival of the direct sound would eliminate one source of interaural cross correlation variability and allow others, relating to the architectural reflections, to become more recognizable.

Results show that removing the direct sound from the interaural cross correlation calculation typically decreases both within-room and among-room variability (refer to figures 33 & 34). Results also show that removing the direct sound from the interaural cross correlation calculation typically decreases the resulting value significantly. As frequency increases, so does the effect of the direct sound. This means that excluding the direct sound from a 4kHz IACC calculation generally decreases the value more than excluding it from a 500Hz calculation. As one would expect, the direct sound affects interaural cross correlation values of front positions in rooms more than the values of rear positions. Therefore, excluding the direct sound from the signals of a seat by the stage decreases the resulting value more than excluding it from the signals of a seat in the rear of the hall.

## Unexplained Variation

Although the relative level of the direct sound accounted for a large part of interaural cross correlation variability in some rooms and at some frequencies, a significant amount of variability still remained unexplained. Further data analysis attempted to identify other sources of variability that related to the architectural design of the rooms and that could not be calculated using a single monaural recording.

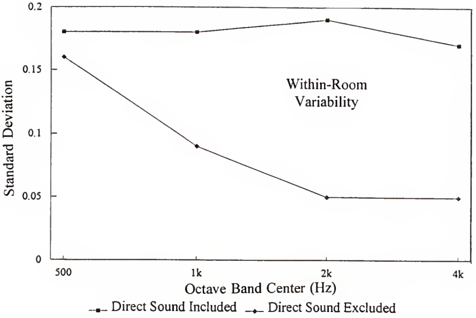



Figure 33 Decrease in standard deviation due to excluding the direct sound

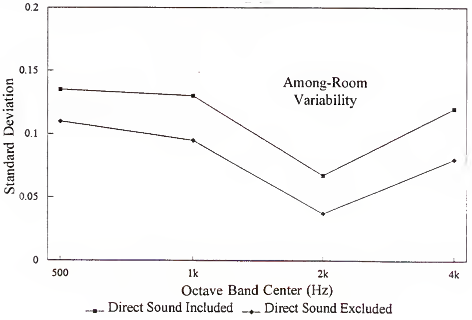
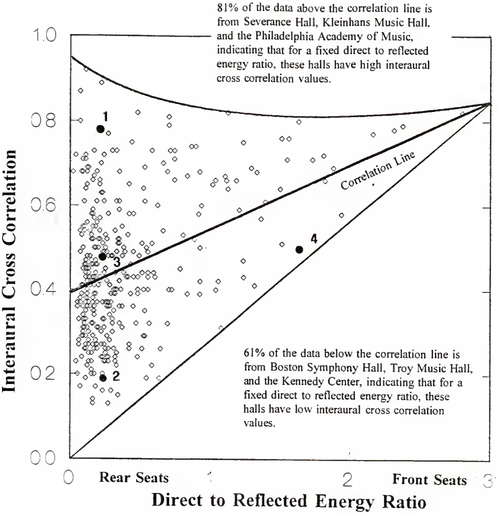



Figure 34 Decrease in standard deviation due to excluding the direct sound

When a scatter plot, such as the one in figure 35, containing all data from all rooms was examined, it was found that the data above the correlation line was typically from one of three concert halls: Severance Hall, Kleinhans Music Hall, or The Philadelphia Academy of Music. Similarly, the data below the line was typically from one of three different concert halls: Boston Symphony Hall, the Kennedy Center, or Troy Music Hall. Architectural comparisons of the halls in both groups began to support the premise that interaural cross correlation relates to the arrival direction of the architectural reflections.

Boston Symphony Hall, The Kennedy Center, and Troy Music Hall are all rectangular halls with solid side walls. It seemed possible that reflections arriving from the sides in these rooms caused the corresponding interaural cross correlation values to fall below the linear correlation line (*i.e.*, to have low IACC values for a given D/R energy ratio). Conversely, it seemed possible that Severance Hall, Kleinhans Music Hall, and The Philadelphia Academy of Music lacked reflections from the sides. As a result, the data from these three halls rose above the linear correlation line (*i.e.*, to have high IACC values for a given D/R energy ratio). Visual inspection supported that the binaural impulse responses from the three former halls had greater interaural time and level differences than those from the three latter halls, but an explanation of the differences remained undetermined.

The architectural reasons that the latter three halls lack reflections from the sides is different for each hall . Kleinhans Music Hall is fan-shaped with a low ceiling relative to those in the rectangular halls. The close proximity of overhead surfaces and the remoteness of the splayed side walls, probably prevents reflections from the sides at many seating locations. The main floor seating area in Severance Hall is surrounded by a deep promenade. Since only a colonnade separates the main floor seating from the promenade (*i.e.*, side walls are not present), the possibility of reflections from the sides is remote. The Philadelphia Academy of Music, although currently used for orchestral performances, was originally designed as a




*Generally, seats in the rear of the concert halls have low direct to reflected energy ratios. As distance from the source decreases, the direct to reflected energy ratio typically increases, so that seats by the stage have the highest direct to reflected energy ratios. As the direct to reflected energy ratio increases, the variation of interaural cross correlation decreases.*

Figure 35 Interaural cross correlation values versus direct to reflected energy ratios

horseshoe-shaped opera house with four steep balconies wrapping completely around the house. The plush seating in these balconies is steeply sloped and continues far underneath the balcony overhead. Again, reflective side walls are not present.

Data point *1* in figure 35 was measured along the central axis inside Kleinhans Music Hall. Since the position is towards the rear of the of the hall, it has a low D/R energy ratio. Since the hall is fan-shaped and there is a lack of reflections from the sides, the IACC value is high. Data point *2* was measured along the central axis inside the Kennedy Center. Since the position is towards the rear of the hall, it too has a low D/R energy ratio. However, since the hall is rectangular and side wall reflections exist, the IACC value is low. Data point *3* was measured along the central axis in the rear of The Philadelphia Academy of Music. Data points *1*, *2*, and *3* all have approximately equal D/R energy ratios, yet different IACC values. Data point *3* has an IACC value that is slightly higher than average when compared to other data points with the same D/R energy ratio. This indicates that data point *3* most likely has neither an abundance of nor a lack of reflections from the sides. Data point *4* was measured along the central axis in the front of Boston Symphony Hall. Even though the IACC values for data points *3* and *4* are equal, information about the relative presence or absence of reflections from the sides is obscured by the varying D/R energy ratios. The IACC value for data point *3* is only average for its D/R energy ratio. Conversely, data point *4* has the lowest IACC value of all data points with similar D/R energy ratios. Therefore, despite the equal IACC values, it is speculated that data point 4 has more reflections approaching from the sides than data point *3*.

Although the amount of interaural cross correlation variability unexplained by the direct to reflected energy ratio seemed to relate to the arrival direction of the architectural reflections, the support thus far was somewhat speculative. After all, there could have been numerous other common factors, architectural or not, relating the concert halls in each of the two groups.

## Discussion

Two sources of interaural cross correlation variability have been identified, and a better understanding of how they interact has been achieved. The first source, the level of the direct sound relative to the level of the early architectural reflections, is calculated using a monaural signal, and leads to no useful binaural information about the architectural reflections. When this source of variability is removed by excluding the direct sound from the interaural cross correlation calculation, both within-room and among-room variability decreases. The second source of interaural cross correlation variability seems to be the architectural design of the room. For a fixed direct to reflected energy ratio, higher interaural cross correlation values seem to relate to rooms that lack side wall reflections (such as fan-shaped rooms). Conversely, lower interaural cross correlation values seem to relate to rectangular rooms with reflective side walls.

However, the interaction of these two sources of interaural cross correlation variability changes with receiver position. Front positions in concert halls consistently have high interaural cross correlation values because the direct sound, which is equal to both ears, is so much louder than the subsequent architectural reflections. These positions will have high interaural cross correlation values regardless of the arrival direction of the succeeding architectural reflections. Therefore, little if any information about the architectural enclosure, beyond that which can be obtained from a monaural signal, can result. Receiver positions in the middle of concert halls have interaural cross correlation values that vary partly due to the level of the direct sound and partly due to the arrival direction of the architectural reflections. The significance of each of these factors on the overall interaural cross correlation value can not be determined using the current data. It seems however, that variability of interaural cross correlation values for rear positions in halls is determined to a large extent by the arrival direction of the architectural reflections. The attenuated direct sound has minimal effect.

# INTERAURAL CROSS CORRELATION EXPERIMENTS IN SCALE MODELS

## Introduction

A 1:10 scale interaural cross correlation measurement system was developed for two purposes, 1) to gain further support for the conclusions that resulted from the real room data analysis and 2) to continue investigating the effect of architectural design on interaural cross correlation values.

## Measurement Method

### Sound Source

The sound source used in the acoustical model experiments was a Grozier Technical Systems electric spark generator designed specifically for ultrasonic modeling. It is impulsive, loud, and has sufficient energy in the desired 1:10 scale bandwidth (2kHz to 50kHz). The Grozier spark is highly repeatable. Multiple trials for one source/receiver combination produced consistently identical interaural cross correlation values. When the multiple shots were averaged and the 95% confidence intervals were calculated, the confidence intervals were so small that variability due to the model measurement system was judged insignificant.

### Binaural Receiver

The design of the 1:10 scale binaural receiver began with the a review of previous work by other researchers. Xiang (1991) recently constructed a manikin from silicon rubber and inserted 1/8" Bruel & Kjaer condenser microphones from the underside of the torso. The tips

of the microphones were inserted far enough to extend into the head and couple with the ear canals (0.8mm inside diameter). A diagram of Xiang's design is shown in figure 36.

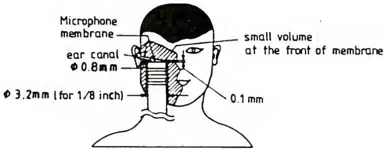



Figure 36 Scale (1:10) manikin design (Xiang, 1991)

Xiang went to remarkable extremes to accurately reproduce the pinnae for the manikin. A full scale pinna replica was sliced into 1mm thick layers. Each of the layers were photographically reduced by a factor of ten and cut from a piece of foil. The reduced layers were then reassembled and used to make a casting mold. The final scaled pinnae were then cast with elastic silicon rubber.

Having previously shown that interaural cross correlation is not sensitive to the filtering effects of the pinnae and auditory canals (refer to the full scale *Measurement Method* section), the simplified 1:10 scale manikin was designed without pinnae and auditory canals. If future research merits collection of scale model impulse responses inclusive of the filtering effects of the pinnae and auditory canals, simple geometric representations of the pinnae can be added to the head of the manikin at that time (Teranishi and Shaw, 1968).

When constructing the full scale manikin, an existing head replica was altered by removing the crown so that instrumentation could be placed inside. This approach did not seem feasible for the construction of the scale manikin. The extremely small size would prevent any work within the head. Instead a mold, commonly used to cast porcelain dolls, was used. The mold

was actually for a 1:12 scale porcelain doll. However, the porcelain greenware normally shrinks when it is fired. Therefore, the mold is actually larger than the 1:12 size. As long as a nonshrinking material is used in lieu of the porcelain, the resulting cast is exactly 1:10 scale. The head and torso of the scale manikin were cast with a clear casting resin (refer to figure 37). The resulting dimensions are compared to those of other scale manikins in table 8.

Figure 37 Scale (1:10) manikin

Table 8 Scale manikin characteristics

| Measurement (mm) | Burkhard (1975) | Genuit | Xiang (1991) | Madaras |
|---|---|---|---|---|
| Head Breadth | 15.1 | 17.7 | 15.5 | 16.9 |
| Head Depth | 18.8 | 21.8 | 21.0 | 20.0 |
| Head Height | 22.2 | 26.1 | 23.3 | 25.0 |
| Ear Canal to Top of Head | 13.0 | 15.6 | 13.4 | 15.5 |
| Ear Canal to Back of Head | 9.0 | 11.6 | 11.6 | 11.0 |
| Ear Canal to Face | 11.2 | 10.4 | 11.9 | 11.0 |
| Neck Breadth | 11.2 | 11.7 | 11.6 | 13.0 |
| Neck Depth | 29.1 | - | 29.2 | 29.0 |
| Chest Breadth | - | 26.9 | 21.0 | 24.0 |
| Chest Depth | - | 26.9 | 21.0 | 24.0 |

Instrumentation, Filtering, and Processing

The scale manikin is equipped with Knowles Electronics EK3132 microphones that measure 5mm x 4mm x 2mm. The signals are amplified by Tucker-Davis Technologies MA2 microphone amplifiers before being acquired (with 12 bit resolution) with a LeCroy 6810 multi-channel digitizer (at a sampling rate of 500kHz). The signal to noise ratio of the scale manikin system (wideband, 50dB; 125Hz octave band, 46dB; 250Hz, 500Hz, 1kHz, & 2kHz octave bands, $62^{+}$dB; 4kHz octave band, 40dB) is better than that of the real room system and also better than that of other scale manikins (Xiang reports a signal to noise ratio of only 30dB). The good signal to noise ratio is attributed to the fact that the microphones are mounted so that their faces are flush with the outer surface of the head, and not buried inside the head and preceded by small auditory canals.

The microphones were primarily chosen for their size and sensitivity, so their frequency response is less than ideal. Between 250Hz and 2kHz (full-size equivalent frequencies) the frequency response is relatively flat (less than 5dB change as shown if figure 38) compared to the frequency response of the gunshot (refer to figure 12). Above 2kHz the frequency response rolls off steeply (20dB down at 4kHz and 40dB down at 8kHz). The left ear and right ear microphones are similar in sensitivity and frequency response (less than 1dB difference between 125Hz and 3kHz).

The ARIAS system, used to process and filter the full scale binaural impulse responses, was also used to filter and process the scale model binaural impulse responses. Before processing, the sampling frequency in the file header was manually changed from 500kHz (1:10 scale model sampling rate) to 50kHz (full scale sampling rate), automatically performing the appropriate scaling.

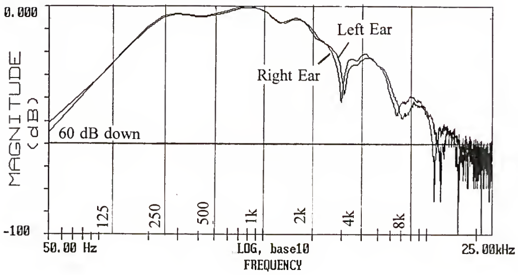



Figure 38 Frequency response of scale measurement system

Scale Model

The model was designed with specific characteristics in an attempt to simplify the study and to make the results easier to interpret (refer to figure 39 and table 9). The first criterion used while designing the model was the necessity for all of the boundary surfaces to be specular reflectors. Therefore, surfaces such as the ceiling, house walls, and stage enclosure are planar, smooth, and hard. Starting with planar surfaces allowed the subsequent addition of various other surface treatments and architectural elements.

The other criterion used while designing the model was to keep the overall room size small enough so that the earliest reflected sound energy would arrive at the receiver positions soon after the direct sound, but large enough to maintain significant within-room variability of interaural cross correlation values. Seats in large rooms, such as the Kennedy Center, often have long initial time delay gaps that result from architectural surfaces being too far from the seated listeners. This gap between the arrival of the direct sound and the early architectural reflections can be as long as 50ms, and does not contain any sound except for ambient noise.

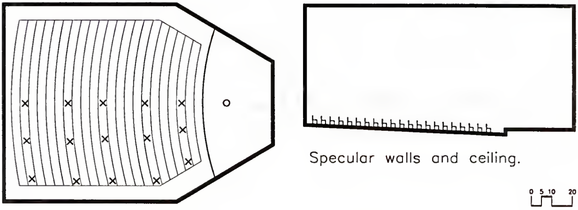



Figure 39 Basic scale model design

Table 9 Concert hall / modeled room size comparison

| | Main Floor Seats | Seating Area | No. of Rows | House Width | House Length | House Height | House Volume | Stage Width | Stage Depth | Stage Area |
|---|---|---|---|---|---|---|---|---|---|---|
| Kennedy | 1,638 | 8,417sf | 34 | 90’ | 125’ | 58’ | 763,501cf | 75’ | 39’ | 2,630sf |
| Boston | 1,456 | 6,241sf | 33 | 75’ | 130’ | 61’ | 670,579cf | 58’ | 33’ | 1,655sf |
| Troy | 847 | 3,680sf | 25 | 70’ | 90’ | 61’ | 399,778cf | 70’ | 23’ | 1,296sf |
| Model | 857 | 5,100sf | 21 | 80’ | 85’ | 60’ | 425,910cf | 48’ | 20’ | 760sf |

Figure 40 shows three left-ear impulse responses measured at different distances from the source along the central axis of a real concert hall (Kennedy Center). Figure 40a shows how positions in the front of the room typically experience long initial time delay gaps and high interaural cross correlation values. Figure 40c shows how positions in the rear of the rooms typically experience short initial time delay gaps and low interaural cross correlation values. This proposed relation seems logical when the length of the initial time delay gap is compared to the length of the integral duration in the interaural cross correlation formula. If $t_2$ in the interaural cross correlation formula (equation 5) is 80ms, and the initial time delay gap extends for 50ms, over half of the integral duration contains no architectural reflections.

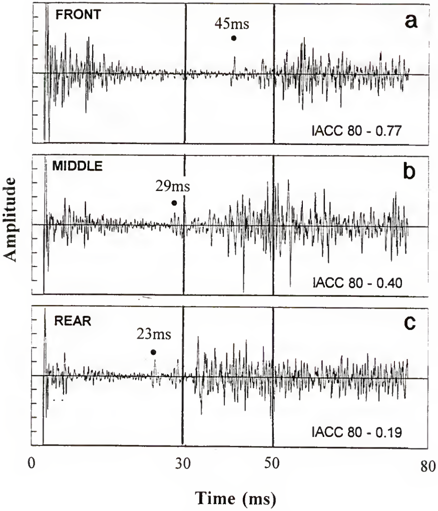



Figure 40 Relationship between interaural cross correlation and initial time delay gap

Therefore, it was believed that the initial time delay gap could possibly be a source of variability among interaural cross correlation values measured in concert halls. Since the model experiments were intended to further investigate the effects of reflection arrival direction on interaural cross correlation, the removal of the variability due to the initial time delay gap was judged sagacious. To eliminate this variability, the proposed research was conducted in a model of a small room so that the initial time delay gap was insignificant (in duration) at all receiver positions.

However, designing a large enough model to maintain a significant amount of within-room variability was also a concern. The modeled room was not as big as some of the larger rectangular concert halls such as the Kennedy Center and Boston Symphony Hall. However, it was similar in size and shape to Troy Music Hall which, despite its small size, does have a significant amount of within-room variability (refer to figure 41).

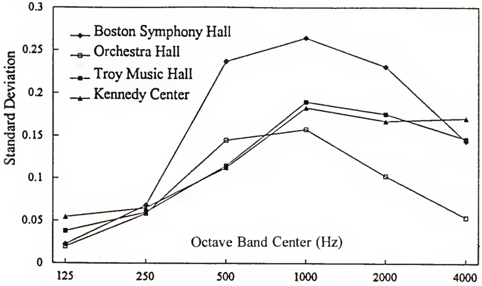



Figure 41 Standard deviation of interaural cross correlation values measured inside four concert halls

Basically, the modeled room was similar in height and width to all three concert halls, however it was shorter in length and had less volume than either the Kennedy Center or Boston Symphony Hall. Except for the stage area, the modeled room was very similar in overall size to Troy Music Hall. The modeled room was designed intentionally with a small stage. It was believed that the smaller stage enclosure would help to decrease the initial time delay gap throughout the room.

The model was designed so that it could be architecturally altered into six different configurations. A description of each is given below. Floor plans and sections showing the differences between the model configurations appear in appendix *B*.

1. *No Hall* Stage floor and house floor with seats only.
2. *Specular* Stage and house enclosed with large flat planar surfaces forming a basic rectangular solid.
3. *Diffusive* Perimeter of the stage and house were treated with a diffusive finish beginning at the floor and extending twenty feet (in scale) up the walls.
4. *Standard* A stage canopy and two tiers of balconies along the side walls and back wall were added to the *specular* configuration. No diffusive finish.
5. *Side* Side balconies were in place. Back balconies and stage canopy were removed. Absorbent material was on the stage back wall, ceiling, and house back wall.
6. *Top/Front/ Back* Side balconies were removed. Back balconies and stage canopy were installed. Absorbent material was placed on the side walls from floor to ceiling.

## Direct to Reflected Energy Ratio

Interaural cross correlation values and direct to reflected energy ratios were measured at multiple receiver positions using the various model configurations. Similar to the results of real room data analysis, a strong relationship between these two parameters was also found in the model (refer to figure 42). Additionally, the strength of this relationship varied among hall configurations and frequency bands in ways that were similar to those evidenced during real room data analysis.

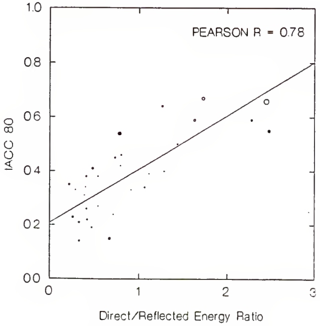



Figure 42 Relationship between model interaural cross correlation values and direct to reflected energy ratios

## Effect of Reflections from the Sides on Interaural Cross Correlation

The model was first used to test the effect of reflection arrival direction on interaural cross correlation. Data from the *side* and *top/front/back* model configurations were correlated with direct to reflected energy ratios. The results showed, that for a fixed direct to reflected energy ratio, the *side* model configuration generally produced lower interaural cross correlation values than the *top/front/back* model configuration (refer to figure 43). Therefore, the speculation that Kleinhans Music Hall, Severance Hall, and the Philadelphia Academy of Music all have less reflections arriving from the sides than Boston Symphony Hall, Troy Music Hall, and the Kennedy Center was supported.

Additionally, there is an area in figure 43 where data from the *side* and *top/front/back* model configurations intermix. These data are almost entirely from seats in the front half of the model, and are more strongly related to the direct to reflected energy ratio than the rest of the data. The intermixing of these data points shows that the interaural cross correlation values for front seats are primarily determined by the level of the direct sound relative to the level of the architectural reflections. After all, two greatly different model configurations produced similar interaural cross correlation values.

There are two other clusters of data in figure 43, one in the upper left corner (*top/front/back* configuration) and the other in the lower left corner (*side* configuration). The data in these two clusters are from the seats in the back half of the model. The tight grouping of the data in these clusters shows that the interaural cross correlation values for seats in the back half of the model are less influenced by the level of the direct sound. The clear separation of the two clusters and the large distance between them shows that seats in the back half of the model are greatly influenced by the arrival direction of the architectural reflections.

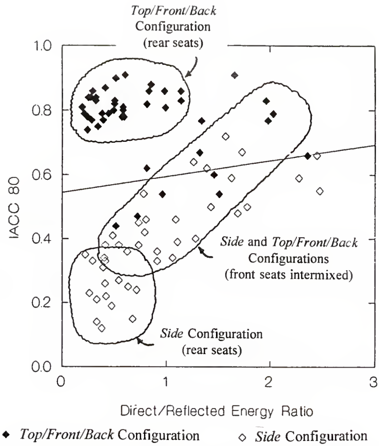



Figure 43 Interaural cross correlation values plotted versus direct to reflected energy ratios for two model configurations

## Effect of Other Architectural Changes on Interaural Cross Correlation

Interaural cross correlation data was measured for all six model configurations. Figure 44 shows the resulting hall averages. Since the measurement system is not a source of variability, even slight differences were attributed to the architectural characteristics being studied.

### *No Hall* versus Others

Since all of the enclosed model configurations resulted in lower interaural cross correlation values than the *no hall* configuration, it can be concluded that architectural reflections in general decrease interaural cross correlation more than the absence of architectural reflections.

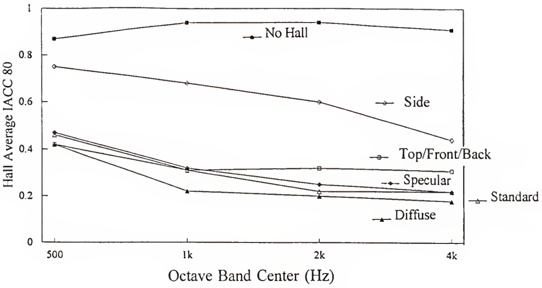



Figure 44 Comparison of interaural cross correlation values measured inside six different model configurations

<u>*Side* versus *Top/Front/Back*</u>

Since the *side* configuration produced values lower than those produced by the *top/front/back* configuration, it can be concluded that interaural cross correlation is sensitive to the arrival direction of the architectural reflections.

<u>*Specular* versus *Absorptive*</u>

One may have expected that adding absorption on the ceiling and back walls of the stage and house would decrease interaural cross correlation relative to an otherwise specular reflecting room. However, the results of this study show that the addition of a large amount of absorption, regardless of its location, increases interaural cross correlation. It is speculated that the added absorption prevented many multiple-order reflections that tend to decrease interaural cross correlation values. It is interesting though, that adding the absorption to the side walls increased interaural cross correlation greatly while adding it to the ceiling and back walls of the stage and house increased interaural cross correlation only slightly.

*Specular* versus *Standard* and *Diffusive*

There were only slight differences between hall averages for the *specular*, *standard*, and *diffuse* configurations. However, the magnitude of these differences are more representational of the differences that occur between real room hall averages. For example, the difference between the hall averages of Boston Symphony Hall (rectangular) and Kleinhans Music Hall (fan-shaped) are not as great as the differences between the hall averages of the *side* and *top/front/back* model configurations. Therefore, even slight changes in interaural cross correlation hall averages can be significant.

The addition of architectural features such as a stage canopy and balconies decreased the hall average slightly relative to the *specular* conditions. It is speculated that the decrease is partly due to the addition of multiple-order reflections arriving from the sides, and partly due to the naturally diffusive edges of the added elements. The addition of diffusive material on the walls decreased the hall average significantly compared to the *specular* and *standard* model configurations. It is expected that adding diffusive materials on the stage canopy and balconies would produce an even lower hall average.

Lastly, it is important to note that the values in figure 44 were calculated with the direct sound excluded. When hall averages that included the direct sound were graphed in a similar manner, the *specular* and *standard* hall averages overlapped, and the *diffusive* hall average actually exceeded the other two hall averages. It was concluded that removing the direct sound from the interaural cross correlation calculation allows more subtle differences in the architectural reflection patterns to have greater effect on the resulting interaural cross correlation values. This means that it was not until after the removal of the direct sound that some of the more subtle differences between the model configurations affected the interaural cross correlation values.

# CONCLUSIONS

## Measurement of Interaural Cross Correlation

This research has measured interaural cross correlation values in multiple concert halls using two different microphone locations relative to the manikin head and a number of binaural receivers with varying levels of detail. Interaural cross correlation values measured with the microphones inside the manikin head (truncating the 2.20cm long auditory canals) were statistically equal to those measured with the microphones located outside the manikin head (approximately 30mm from the ear canal openings). The variability of the interaural cross correlation measurement system was not a factor during this part of the study since both inside and outside impulse responses resulted from the same gunshot. Interaural cross correlation values measured using multiple binaural receivers with increasingly simplified detail in the face and ears were all approximately the same as those measured using a highly detailed manikin. The slight differences among the varying levels of manikin detail fell within the variability of the interaural cross correlation measurement system. In order to establish whether the slight interaural cross correlation differences were attributed to the actual level of detail in the manikin and not the variability of the measurement system, the experiment needs to be performed again with a more repeatable electronic source.

Since a detailed binaural hearing manikin with accurate pinnae and auditory canals is not required for the measurement of interaural cross correlation, more simplified and cost effective representations of the human body can be used by a larger number of researchers to continue studying the variability of interaural cross correlation in concert halls. The ability to use a

more simplified binaural receiver has a direct application to studying interaural cross correlation variability inside scale models. This new method of studying interaural cross correlation should greatly facilitate and expedite future research that attempts to relate interaural cross correlation and architectural features of concert halls.

## Interaural Cross Correlation: Sources of Variability in Concert Halls

Analysis of data from multiple real rooms and scale models has identified three primary sources of interaural cross correlation variability in concert halls. The first source of variability, which can account for as much as half of the interaural cross correlation variability in concert halls, is the level of the direct sound relative to the level of the early arriving architectural reflections. Seats located near the stage of a concert hall generally experience a direct sound that is much higher in level than the succeeding architectural reflections. The direct sound reaches the two ears simultaneously with equally great amplitude. As a result, seats near the stage typically have higher interaural cross correlation values because the direct sound proportionately dominates the overall amount of energy that reaches the receiver.

Seats in the rear of the room generally experience a direct sound that is much lower in level relative to the succeeding architectural reflections. Even though the direct sound reaches the two ears simultaneously, it has been greatly reduced by factors such as air absorption, natural geometric spreading, and attenuation from grazing over seat backs and diffracting over balcony fasciae. As a result, the direct sound has no more influence on the overall interaural cross correlation value than a typical first order reflection off the stage enclosure, ceiling, or walls. Therefore, seats in the rear of concert halls usually have lower interaural cross correlation values. However, the relative level of the direct sound is calculated using a monaural signal, and therefore it yields no useful binaural information relating to the architectural features of concert halls.

The second source of interaural cross correlation variability, which accounts for a significant part of the variability remaining unexplained by the relative level of the direct sound, is the general arrival direction of the early architectural reflections. This directional characteristic can not be measured using a monaural recording and does in fact produce binaural information that relates to the architectural features of concert halls. Reflections approaching the receiver from the sides, such as those off solid side walls in narrow rectangular concert halls, result in lower interaural cross correlation values. Conversely, reflections approaching from above and behind the receiver, such as those off low ceilings in fan-shaped concert halls, result in higher interaural cross correlation values.

The independent significance of each of these two sources of variability on the overall interaural cross correlation value varies with seat location and can not be easily determined with the data in this study. Generally though, interaural cross correlation values for seats near the stage are greatly influenced by the relative level of the direct sound. The architectural reflections, regardless of their arrival direction, are down too far in level relative to the direct sound to effect the overall interaural cross correlation value. As a result of being highly influenced by the relative level of the direct sound and not the arrival direction of the reflections, seats near the stage have consistently high interaural cross correlation values.

Conversely, interaural cross correlation values for seats in the rear of concert halls are not significantly affected by the relative level of the direct sound. Instead, the overall interaural cross correlation value is primarily dependent on the general arrival direction of the early architectural reflections. As a result, architectural reflections arriving from the sides of the receiver produce significantly lower interaural cross correlation values than do reflections approaching from above and behind the receiver. Interaural cross correlation values for seats in the center of concert halls vary partly due to the relative level of the direct sound and partly

due to the general arrival direction of the early architectural reflections. Therefore, interaural cross correlation does relate to the arrival direction of architectural reflections as long as the relative level of the direct sound is low enough, such as in the rear seats of concert halls, to have minimal effect on the overall value.

Because the significance of these two types of variability on the overall interaural cross correlation value varies, seat to seat comparisons of interaural cross correlation values can result in unreliable information about the arrival direction of the architectural reflections. For example, assume that two seats in a small hall, one near the stage and one in the rear of the room, have equal interaural cross correlation values. This does not necessarily mean that the directional characteristics of the sound field at those two positions are equal. Most likely, the seat near the stage has more sound arriving from the sides since it undoubtedly has a higher direct to reflected energy ratio. Similarly, an equal interaural cross correlation value for the same seat in two different concert halls does not necessarily mean that the arrival direction of the architectural reflections for both seats is equal, for the two seats could still have different direct to reflected energy ratios. One method of making with-in room or among-room seat to seat comparisons of interaural cross correlation values produce reliable information about the arrival direction of architectural reflections is to compare only seats having equal, or at least similar, direct to reflected energy ratios. This way, for a given direct to reflected energy ratio, higher interaural cross correlation values should indicate that architectural reflections are arriving from the sides of the receiver. Lower interaural cross correlation values should indicate that architectural reflections are arriving from above and behind the receiver.

However, comparing only seats with equal direct to reflected energy ratios is limiting and makes within-room seat to seat comparisons of interaural cross correlation values impossible. Seats near the stage will always have higher direct to reflected energy ratios than seats in the

rear of the room. Instead, it is suggested that the variability due to the relative level of the direct sound be eliminated by excluding the direct sound from the interaural cross correlation integral durations. This is justified because the relative level of the direct sound is calculated using a monaural signal, and therefore its effect on the variability of interaural cross correlation in concert halls provides no useful binaural information about the architectural reflections.

Since the purpose of this research was primarily to determine if interaural cross correlation could be used to gain binaural information about the arrival direction of architectural reflections, the interaural cross correlation variability due to the relative level of the direct sound was removed by excluding the direct sound from the interaural cross correlation integral duration. As a result of excluding the direct sound from the interaural cross correlation calculation, both within-room and among-room variability of interaural cross correlation values significantly decreased. Once the variability due to the relative level of the direct sound is removed, any seat to seat comparison can be informative, because the remaining variability is then due primarily to the arrival direction of the architectural reflections.

The third source of variability of interaural cross correlation values in concert halls is the placement of smaller architectural elements (such as a stage canopy or balconies) within the room and the finish of the architectural surfaces. Although the effect of this source of variability on interaural cross correlation values is smaller than that resulting from the other two sources of variability, the placement of smaller elements within the room and the finish of the surfaces must be considered in order to achieve the lowest interaural cross correlation values.

Continued research is needed to investigate the effect of more specific architectural features such as stage canopy height and design or balcony depth and height on interaural

cross correlation. However, this research project showed that in order to achieve low interaural cross correlation values, several architectural guidelines should be followed. First, and most important, architectural reflections should approach the listeners from the sides. This can be achieved by keeping a narrow average room width and placing architectural surfaces so that their orientation and proximity to the audience will result in reflections that approach the listeners from the sides within the first 80ms after the direct sound.

However, reflection arrival direction can only lower interaural cross correlation values to a certain extent. To decrease values even further, the surfaces supplying the reflections should be diffusive. This can be achieved with irregular surface articulations or with the edges of smaller architectural elements (such as balconies) within the room. The lowest interaural cross correlation values (calculated without the direct sound) in real concert halls were found in narrow rooms that also had great amounts of diffusion. It should be noted that these types of rooms actually have higher interaural cross correlation values and greater within-room variability when the direct sound is included in the integral duration. Once the direct sound is excluded from the calculation, there is a great decrease in interaural cross correlation values (especially for seats near the stage), and the within-room variation of interaural cross correlation values becomes insignificant. Lastly, if absorption is added to the room, one should expect a general increase in interaural cross correlation values. However, the increase in interaural cross correlation values can be minimized if the absorption is placed on high ceiling surfaces (in halls with suspended stage reflectors) and house back walls.

It seems then, that interaural cross correlation, as it was proposed by Cremer and Mueller, is still subject to a significant amount of variability unrelated to the binaural directional characteristics of the architectural reflections. If the unrelated variability is removed by excluding the direct sound from the integral duration, the parameter is a much better indicator

of the general arrival direction of architectural reflections in concert halls (*i.e.*, whether the architectural reflections are approaching the receiver from the sides or from above and behind).

The relative importance to listener preference of whether the architectural reflections arrive from the sides or from above and behind has been well documented in the literature, even beyond that which has already ready been reviewed in the earlier sections of this research. Barron (1971) and Barron and Marshall (1981) performed experiments by seating listeners in an anechoic chamber and simultaneously playing music (that had previously been recorded by a chamber orchestra in an anechoic space) through two loudspeakers; one in front of the listeners and another at varying angles to the sides of the listeners. It was concluded that early architectural reflections approaching the listener from the sides are essential for the creation of the qualitatively desired feeling of spaciousness (*i.e.,* a sense of being immersed in the sound that results in part from an apparent widening of the sound source).

Most recently, Soulodre and Bradley (1993) conducted a pilot study using methods similar to those of Barron and Marshall. Their results supported the findings of Barron and Marshall as well as those previously found by Keet (1968), namely that sound fields containing architectural reflections that approach the listener from the sides are judged to have an apparently wider sound source. Later, Soulodre and Bradley (1994) after continued study concluded that spatial impression is composed of at least two parts; apparent source width and listener envelopment. Apparent source width is influenced by early reflections approaching from the sides of the listeners, but is less apparent in the presence of reverberant energy. Listener envelopment is produced by later arriving energy and is effected more by level and arrival time than by arrival direction.

The literature supports that lower interaural cross correlation values strongly relate to listener preference, but does not necessarily provide a complete understanding of how lower interaural cross correlation values can be achieved architecturally in real concert halls. This research has shown with real room data and scale model data that lower interaural cross correlation values can be achieved by designing concert halls so that the majority of reflections approach the listener from the sides. However, this research has also shown that in order for interaural cross correlation to be affected by the arrival directions of architectural reflections, another more influential source of variability, namely the relative level of the direct sound, must first be eliminated. The study ends full circle by having the literature also support that a qualitatively preferred feeling of spaciousness results from architectural reflections that approach the listener from the sides.

## Future Research

The progress made by this research towards understanding the variability of interaural cross correlation in concert halls is only an initial step. Despite attempts by earlier researchers to normalize interaural cross correlation and eliminate sources of variability unrelated to the directional characteristics of concert halls, it has been found that the parameter still varies due to at least one factor that does not produce useful binaural information. However, the effect of this factor on the variability of interaural cross correlation in concert halls can be eliminated quite easily by excluding the direct sound from the integral duration. Future research should investigate in more detail the amount of interaural cross correlation variability still remaining after that which is related to the relative level of the direct sound has been removed. This research has found a general relationship between the arrival direction of architectural reflections and the remaining amount of interaural cross correlation variability using data from both real rooms and models.

Future progress can be made in real concert halls by developing a real room measurement system, perhaps one with and electronic source, that has a smaller amount of trial-to-trial variability. A more consistent measurement system would allow for more detailed study of interaural cross correlation. Additional data should be collected using the new measurement system in more concert halls. Rooms chosen for future studies should be selected carefully to allow for specific architectural comparisons. For example, measurement of interaural cross correlation in five fan-shaped halls and five rectangular halls of similar capacity could produce conclusive results relating the effect of room shape on interaural cross correlation. In the current research, the small number of halls in each of these shape categories somewhat limits the types of studies that can be performed.

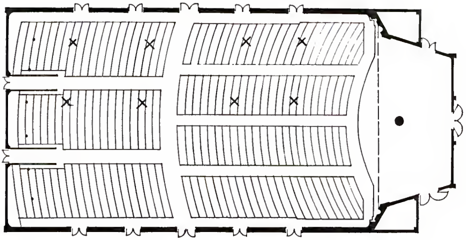
Considering the time and expense of real room measurements, a better way to conduct future interaural cross correlation research may be to continue using scale models. This research used a 1:10 scale model which still proved to be time consuming and costly when major architectural differences were studied. The only factors preventing the study of interaural cross correlation in smaller models is the size, signal to noise ratio, and frequency response of the microphones. Since these factors are purely a technological limitation, perhaps only time needs to pass before research in smaller models is feasible.

Regardless of the scale of the model, the effects of specific architectural changes on interaural cross correlation could be studied. For example, controlled studies varying only one architectural feature such room shape, size, or proportioning (which are not possible in real rooms) could greatly advance the understanding of interaural cross correlation variability in concert halls. This research began to show that fundamental decisions made during schematic design such as room shape have far greater effect on interaural cross correlation values than do the elements placed inside the room such as stage canopies and balconies. Continued research

could investigate the effect or more specific architectural features such as stage canopy height and design or balcony depth and height on interaural cross correlation.

These suggestions for future research are the next natural steps related only to this research. If one steps back and looks at the overall goal of being able to predict qualitative response in a space not yet constructed, many other questions, still related to interaural cross correlation, arise. First, does preference for music listening during actual performances consistently relate to interaural cross correlation values measured during the same performances? Do conclusions based on simulated sound fields using isolated reflections in otherwise anechoic conditions also apply to real sound fields that for the most part contain thousands of reflections in otherwise diffusive conditions? Should interaural cross correlation be minimized, or is there an ideal value or range of values? What is the human threshold for distinguishing interaural cross correlation differences? To what extent must the architecture be changed in order to gain an appreciable decrease in interaural cross correlation. These and many other questions still need to be answered before interaural cross correlation is as useful to room designers as reverberation time. Perhaps in the future when interaural cross correlation has been studied to the extent that reverberation time has today, many of these questions will be answered. This research successfully advanced the current understanding of interaural cross correlation and established a foundation and path for succeeding researchers.

APPENDIX A
PLANS AND SECTIONS OF CONCERT HALLS



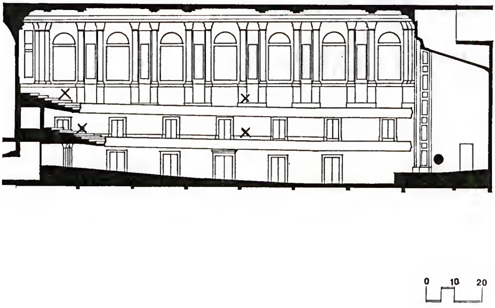



Figure 45 Boston Symphony Hall floor plan and section

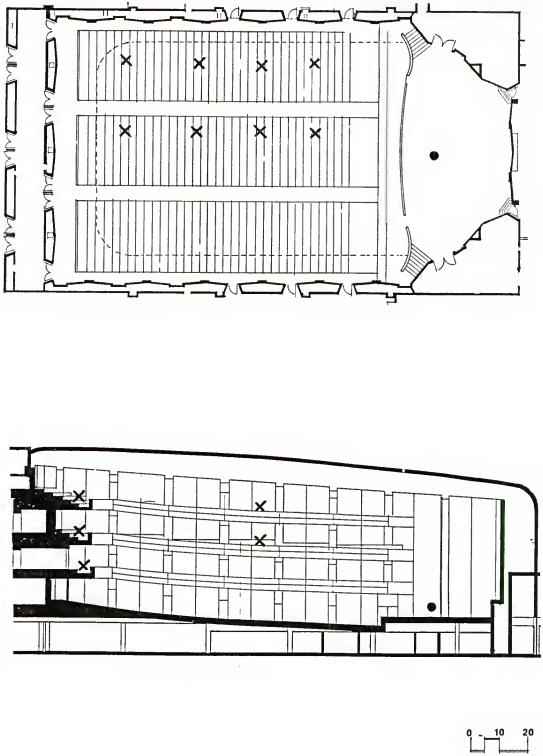



Figure 46 J.F. Kennedy Center floor plan and section

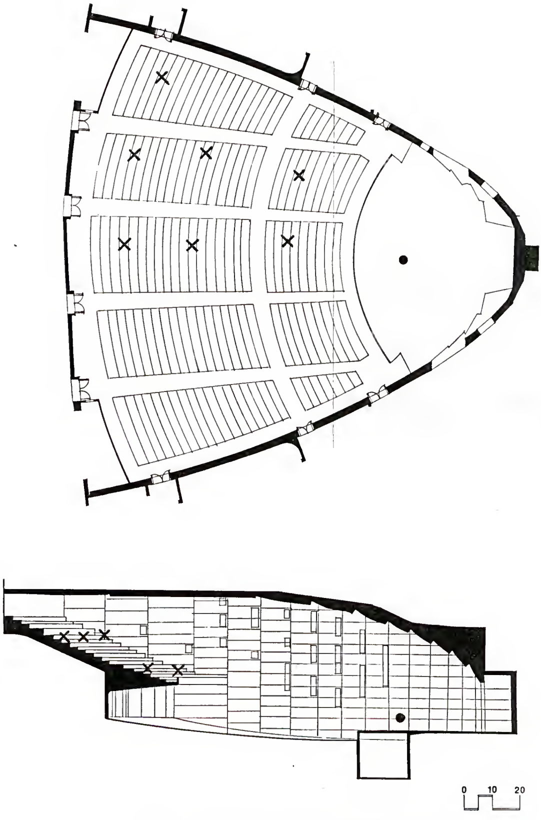
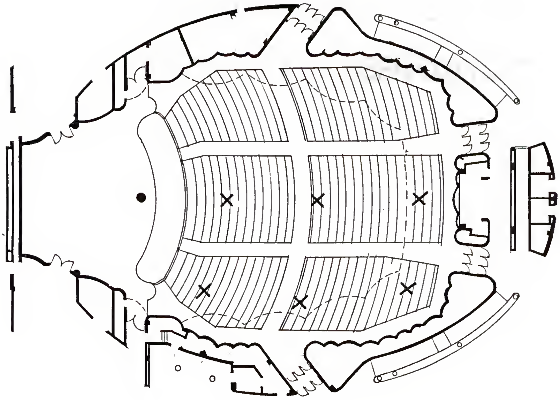



Figure 47 Kleinhans Music Hall floor plan and section



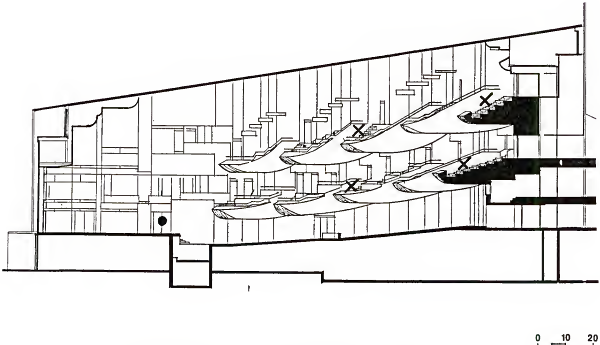



Figure 48 Meyerhoff Concert Hall floor plan and section

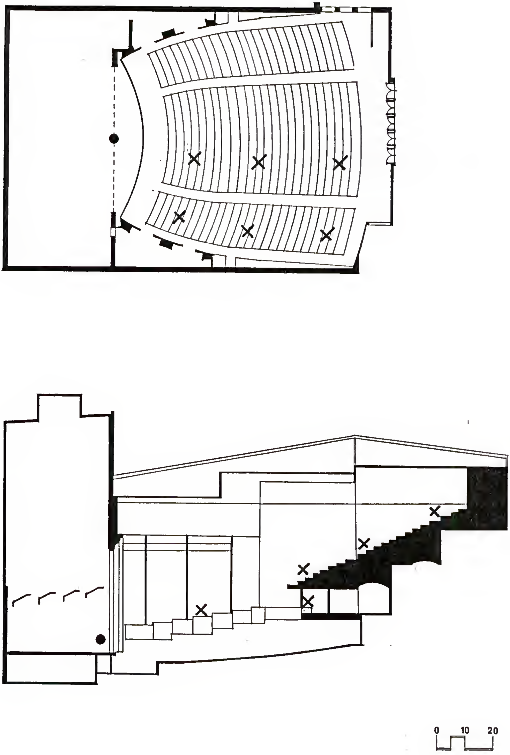



Figure 49 Orchestra Hall floor plan and section

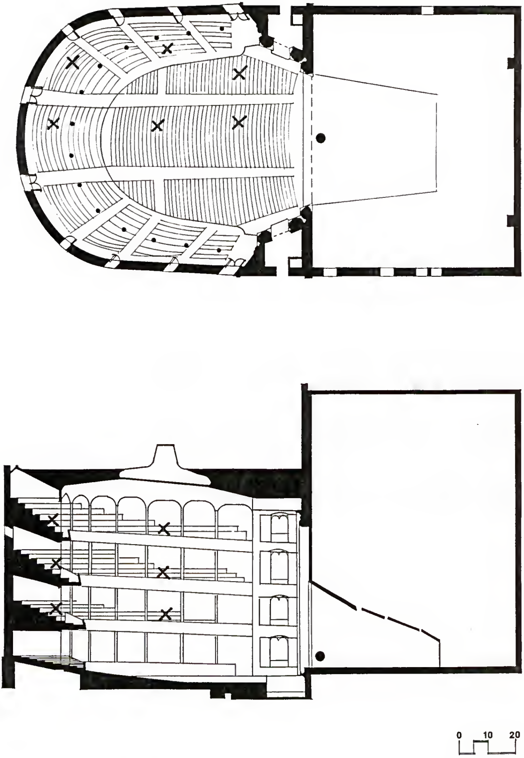



Figure 50 Philadelphia Academy of Music floor plan and section

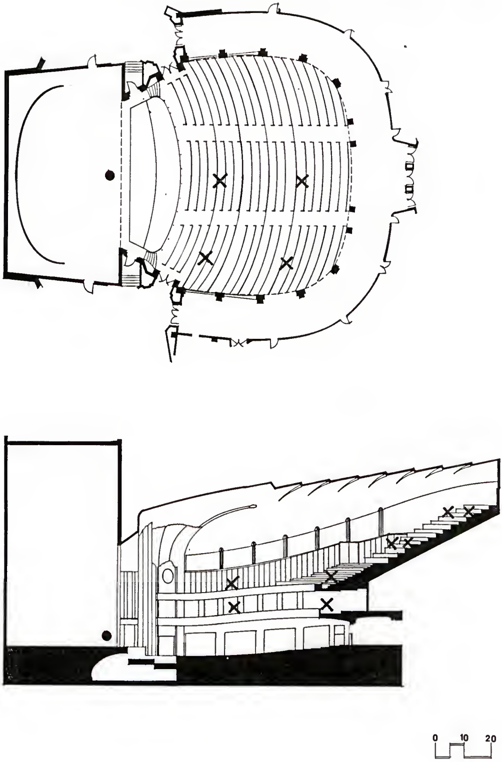



Figure 51 Severance Hall floor plan and section

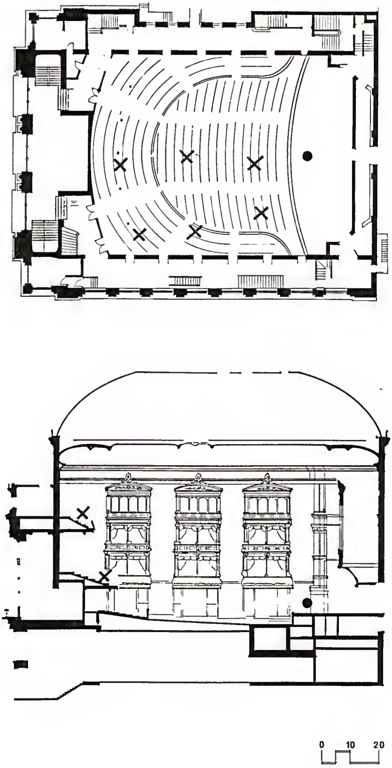



Figure 52 Troy Music Hall floor plan and section

# APPENDIX B
# PLANS AND SECTIONS OF SCALE MODEL CONFIGURATIONS

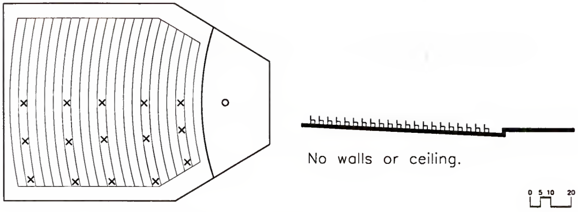



Figure 53 *No hall* model configuration

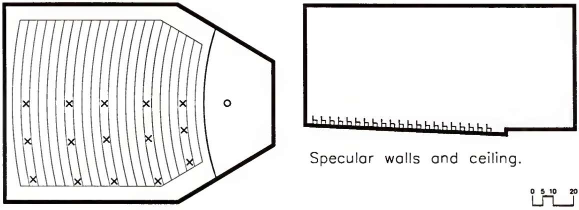



Figure 54 *Specular* model configuration

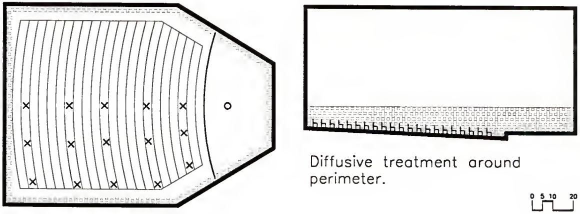



Figure 55 *Diffusive* model configuration

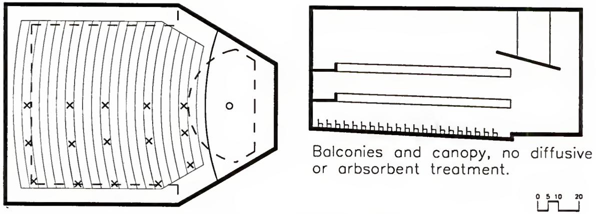



Figure 56 *Standard* model configuration

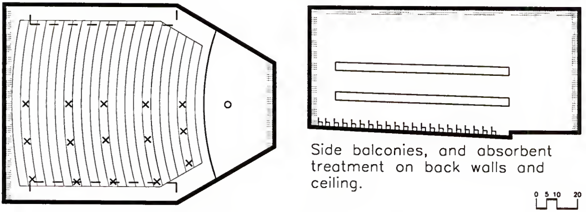



Figure 57 *Side* model configuration

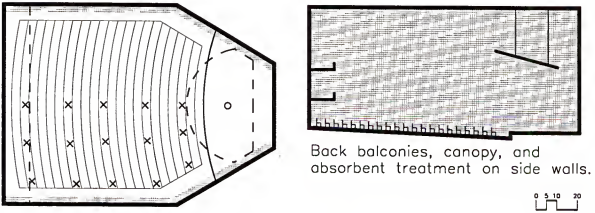



Figure 58 *Top/front/back* model configuration

## LIST OF REFERENCES


Ando,Y. (1985). *Concert Hall Acoustics*, (Springer-Verlag, Berlin).

Ando, Y. (1977b). "Subjective preference in relation to objective parameters of music sound fields with a single echo," J. Acoust. Soc. Am. 62 (6), 1436-1441.

Ando, Y. and Gottlob, D. (1979a). "Effects of early multiple reflections on subjective preference judgments of music sound fields," J. Acoust. Soc. Am. 65 (2), 524-527.

Ando, Y. and Imamura, M. (1979b). "Subjective preference tests for sound fields in concert halls simiulated by the aid of a computer," J. Sound. & Vib. 65 (2), 229-239.

Ando, Y. and Kageyama, K. (1977a). "Subjective preference of sound with a single early reflection," Acustica 37, 111-117.

Ando, Y. and Nakajima, T. (1994). "Effect of IACC on speech clarity and articulation," Internoise 94, Yokohama, Japan, August 29-31.

Barron, M. (1971). "The subjective effects of first reflections in concert halls - The need for lateral reflections," J. Sound Vib. 15, 475-494.

Barron, M. and Marshall, A.H. (1981). "Spatial impression due to early lateral reflections in concert halls," J. Sound Vib. 77, 211-232.

Baur, B., Rosenheck, A., and Abbagnaro, L. (1967). "External ear replica for acoustical testing," J. Acoust. Soc. Am. 42 (1), 204-207.

Beranek, L. (1996). *Concert and opera halls: How they sound*, (Acoustcial Society of America, Woodbury, NY).

Beranek, L. (1992). "Concert Hall Acoustics-1992a," J. Acoust. Soc. Am. 92 (1), 1-39.

Bradley, J. (1986). "Auditorium acoustic measures from pistol shots," J. Acoust. Soc. Am. 80, 199-205.

Burkhard, M. and Genuit, K. (1992). "Artificial head measurement systems for subjective evaluation of sound quality," J. Sound. & Vib. March, 18-23.

Burkhard, M. and Sachs, R. (1975). "Anthropometric manikin for acoustic research," J. Acoust. Soc. Am. 58, 214-222.

Chiang, W. (1994). "Effects of various architectural parameters on six room acoustical measures in auditoria," dissertation, University of Florida.

Cremer, L. and Mueller, H. (1982). *Principles and Applications of Room Acoustics, Vol. 1*, English translation by T. Schultz (Applied Science Publishers, New York).

Damaske, P. (1968). "Subjekive Untersuchug von Scallfeldern," Acustica 19, 199-213.

Hidaka, T., Okano, T., and Beranek, L. (1991). "Studies of interaural cross correlation (IACC) and its relation to subjective evaluation of the acoustics of concert halls," 122nd meeting of Acoustical Society of America, Houston, Texas.

Keet, W. de V. (1968). "The influence of early lateral reflections on spatial impression," 6th International Congress on Acoustics, Tokyo.

Knudsen, E. (1982). "Auditory and visual maps of space in the optic tectum of the owl," J. Neurosci. 2, 1177-1194.

Middlebrooks, J. and Green, D. (1990). "Directional dependence of interaural envelope delays," J. Acoust. Soc. Am. 87 (5), 2149-2162.

Middlebrooks, J., Makous, J., and Green, D. (1989). "Directional sensitivity of sound pressure levels in the human ear canal," J. Acoust. Soc. Am. 86 (1), 89-108.

Morimoto, M. and Iida, K. (1991). "Relation between auditory source width in various sound fields and degree of interaural cross correlation," presented at an international symposium on acoustics, Copenhagen, August 19-22.

Schroeder, M., Gottlob, D., and Siebrasse, K. (1974). "Comparative study of European concert halls: Correlation of subjective preference with geometric and acoustic parameters," J. Acoust. Soc. Am. 56, 1195-1201.

Scroggie, E. (1992). consulting clinical audiologist, Audio Hearing Services, 1905 N.W. 13th St., Suite 1, Gainesville, Florida 32609.

Shaw, E. (1968). "Ear canal pressure generated by a free sound field," J. Acoust. Soc. Am. 39, 465-470.

Soulodre, G. and Bradley, J. (1994). "The influence of late arriving energy on concert hall spatial impression," proceedings of the Wallace Clement Sabine Centennial Symposium in conjunction with the 127th meeting of the Acoustical Society of America, 101-104.

Soulodre, G., Bradley, J., and Popplewell, N. (1993). "Pilot study of simulated spaciousness," presented at the 125th meeting of the Acoustical Society of America, Ottawa.

Teranishi, R. and Shaw, E. (1968). "External ear acoustic models with simple geometry," J. Acoust. Soc. Am. 44, 257-263.

Thiele, R. (1953). "Richtungsverteilung und Zeitfolge der Schallrueckwuerfe in Raumen," Acustica 3 , 291-302.

Xiang, N. (1991). "A mobile universal measuring system for the binaural room acoustic modeling technique," Akustik, Ruhr-Universitat Bochum.

## BIOGRAPHICAL SKETCH

Gary S. Madras was born in 1967, and was raised along with his older sister, Darlene, by his father and mother, William and Ellen Madras, in Amherst, Ohio. He attended Walter G. Nord Jr. High School and Marion L. Steele High School in the Amherst School District. In addition to school, Gary was actively involved in both high school athletics and scouting.

After his high school graduation in the summer of 1985, Gary began his continued studies in the Department of Architecture at Kent State University in Kent, Ohio. From 1985 through 1989 Gary earned his first degree, Bachelor of Science. Gary earned his Bachelor of Architecture degree in the spring of 1991. His fifth year architecture project involved an independent study with Professor Charles Harker, investigating the effect of the computer as a tool on the architectural design process. In the summer of 1991, Gary completed his Master of Architecture degree at Kent State University. His thesis was titled Ideal Reverberation Time Curves for Classical and Romantic Symphonies: An Acoustical Analysis of Severance Hall, Cleveland. It earned the Robert B. Newman Award for Merit in Architectural Acoustics as well as a Student Senate Thesis Award.

As part of his academics, Gary was a member of Kent State's Honor College and served as a teaching assistant in environmental technology and computer graphics. In addition to his studies, Gary worked part-time in two Cleveland architecture firms and was an active member of Delta Tau Delta fraternity.

Gary relocated to Gainesville in the summer of 1991 to begin his doctoral studies in The University of Florida's Department of Architecture under the supervision of Professor Gary W.

Siebein. While earning his Ph.D. Gary served as a teaching assistant in both environmental technology and computer electives, as well as being a research assistant in the architectural acoustics research lab. He is currently employed as an acoustical consultant with Jaffe Holden Scarbrough Acoustics Inc. in Norwalk, CT.

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Earl M. Starnes, Chair
Professor Emeritus of Architecture

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Gary W. Siebein, Cochair
Professor of Architecture

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Bertram Y. Kinzey Jr.
Professor Emeritus of Architecture

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

David M. Green
Graduate Research Professor of Psychology

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Willis R. Bodine Jr.
Professor of Music

This dissertation was submitted to the Graduate Faculty of the College of Architecture and to the Graduate School and was accepted as partial fullfillment of the requirements for the degree of Doctor of Philosophy.

May, 1996

Dean, College of Architecture

Dean, Graduate School